

Query of: US-09-485-529-104 to: GenEmbl:\* out\_format : pfs  
 Date: Dec 19, 2001 11:15 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 CompuGen Ltd.  
 Command line parameters:  
 -MODE=frame+p2n.spool-DEV-xlp  
 -O=/sgn2.1/USPTO.spool/US09485529/runat\_19122001\_165005\_14341/app\_query.fasta\_1.77  
 -DB=GenEmbl -QFMT=fascap -SUFFIX=p2n.rge -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -YGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
 -TRANS=human40.cdi -L1ST=1000 -DOCALIGN=200 -THR\_SCORE=pcpt  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
 -NORM=ext -MINLEN=0 -MAXLEN=200000000  
 -USER=US09485529\_ECGN1\_1\_4773 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
 -WAIT -THREADS=1

Search information block:  
 Query: US-09-485-529-104  
 Query length: 17  
 Database: GenEmbl:\*  
 Database sequences: 1472140  
 Database length: -341344837  
 Search time (sec): 1408.760000

score.list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:AX005896	+	81.00	303.58	1.2e+08	51   AX005896 Sequence 105 from Patent
gb_pat:AX005857	+	81.00	284.25	1.5e-07	453   AX005857 Sequence 66 from Patent
gb_pl:TAE242531	+	81.00	271.70	7.4e-07	1872   A242531 Triticum aestivum rht
gb_pat:AX005805	+	81.00	270.58	8.5e-07	2125   AX005805 Sequence 14 from Patent
gb_pat:AX005794	+	81.00	268.43	1.1e-06	2709   AX005794 Sequence 3 from Patent
gb_pat:AX005807	+	78.00	276.71	3.9e-07	302   AX005807 Sequence 16 from Patent
gb_pat:AX005808	+	78.00	274.89	4.9e-07	371   AX005808 Sequence 17 from Patent
gb_pat:AX005803	+	78.00	268.43	1.1e-06	770   AX005803 Sequence 12 from Patent
gb_pl:ZMA242530	+	78.00	260.48	3.1e-06	1890   A242530 Zea mays partial d8 g
gb_pat:AX005806	+	78.00	258.92	3.8e-06	2255   AX005806 Sequence 15 from Patent
gb_pl:ABX030956	+	78.00	258.01	4.3e-06	2500   ABX030956 Oryza sativa mRNA for
gb_pl:AC087797	+	78.00	223.58	0.0004	122497   AC087797 Oryza sativa chromo
gb_pat:AX005810	+	72.00	246.70	1.8e-05	725   AX005810 Sequence 19 from Patent
gb_pat:AX005867	+	70.00	246.40	1.9e-05	324   AX005867 Sequence 76 from Patent
gb_pl:ATRG62	+	70.00	230.52	0.0001	1951   Y11337 A.thaliana mRNA for RGA
gb_pat:ATRG67	+	70.00	230.46	0.0001	1964   A64697 Sequence 1 from Patent
gb_pl:ATY15193	+	70.00	230.46	0.0001	1964   Y15193 Arabidopsis thaliana CA
gb_pl:ATY15194	+	70.00	229.45	0.0002	2201   Y15194 Arabidopsis thaliana mR
gb_pl:ATRG61	+	70.00	229.42	0.0002	2201   Y11336 A.thaliana mRNA for RGA
gb_pl:ATY2242957	+	70.00	223.99	0.0003	4081   A2224957 Arabidopsis thaliana
gb_pl:ATY3M16	+	70.00	207.40	0.0028	26604   AL391150 Arabidopsis thaliana
gb_pl:ATAC009895	+	70.00	197.41	0.0102	82289   AC009895 Arabidopsis thaliana
gb_pl:AC020665	+	70.00	197.05	0.0106	85702   AC020665 Arabidopsis thaliana
gb_pl:AC005560	+	70.00	196.13	0.0120	95137   AC005560 Arabidopsis thaliana
gb_pl:AC006917	+	70.00	193.18	0.0175	132699   AC006917 genomic sequence fo
gb_pat:AX081276	+	66.00	216.49	0.0009	1779   AX081276 Sequence 1 from Patent
gb_pat:AX081278	+	66.00	216.49	0.0009	1779   AX081278 Sequence 3 from Patent
gb_pat:AX005860	+	65.00	232.12	0.0001	200   AX005860 Sequence 69 from Patent
gb_pat:AX005809	+	62.00	214.51	0.0011	416   AX005809 Sequence 18 from Patent
gb_pl:AF224770	+	51.00	171.65	0.2765	524   AF224770 Andira aubletii clone
gb_htg:AC015886	-	49.00	112.15	570.13	188716   AC015886 Mus musculus clone
gb_pr:AL335796	-	48.00	110.35	718.29	152086   AL335796 Human DNA sequence
gb_pl:AC007504	-	47.00	108.37	925.55	125021   AC007504 Arabidopsis thalian
gb_htg:AC034100	-	47.00	105.17	1.4e+03	179481   AC034100 Mus musculus chromo
gb_ba:AP002977	-	47.00	99.79	2.8e+03	329709   AP002977 Mesorhizobium loti
gb_pat:AX005855	+	46.00	157.77	1.64	309   AX005855 Sequence 64 from Patent
gb_pl:RPX02	+	46.00	96.56	4.2e+03	312430   AJ235221 Rickettsia prowazek
gb_pl:HB040402	+	45.00	143.01	10.89	1078   U040402 Hevea brasiliensis hydr
gb_pat:AX9586	+	45.00	142.90	11.04	1091   AX9586 Sequence 1 from Patent
gb_htg:AC026377	+	45.00	98.68	3.2e+03	161592   AC026377 Mus musculus chromo

gb_htg:AC091618	+	45.00	96.17	4.4e+03	214605   AC091618 Rattus norvegicu
gb_htg:AC016814	+	45.00	95.63	4.7e+03	228201   AC016814 Mus musculus clo
gb_htg:AC020887	+	45.00	94.09	5.8e+03	271519   AC020887 Mus musculus clo
gb_htg:AC070284	+	44.00	141.11	13.89	878   AF070284 Giardia intestinali
gb_pr:AF090137	+	44.00	138.05	20.57	1241   AF090137 Homo sapiens carbo
gb_pat:EL6306	+	44.00	136.63	24.70	1458   EL6306 cDNA encoding kerat
gb_ba:BSORF01	+	44.00	133.41	37.29	2096   X98341 B.subtilis orfRM1 DN
gb_pr:HSUG5637	+	44.00	133.03	39.19	2190   U65637 Homo sapiens chondro
gb_pr:AB003791	+	44.00	132.16	43.79	2415   AB003791 Homo sapiens mRNA
gb_vi:AF298771	+	44.00	126.90	85.92	4374   AF298771 Porcine reproducti
gb_vi:AF298771	+	44.00	123.08	140.32	6739   AF066384 Porcine reproducti
gb_vi:AF066384	+	44.00	118.07	266.74	11870   AF033356 Porcine reproduct
gb_vi:AF033356	+	44.00	118.06	267.05	11882   AF033356 Porcine reproduct
gb_vi:AF033357	+	44.00	118.06	267.05	11882   AF033357 Porcine reproduct
gb_vi:AF159149	+	44.00	115.77	358.16	15390   AF159149 Porcine reproduct
gb_vi:AF325691	+	44.00	115.77	358.16	15390   AF325691 Porcine reproduct
gb_vi:PRU87392	+	44.00	115.76	358.72	15411   U87392 Porcine reproductiv
gb_vi:AF066183	+	44.00	115.76	358.74	15412   AF066183 Porcine reproduct
gb_vi:AF046869	+	44.00	115.75	359.16	15428   AF046869 Porcine reproduct
gb_vi:AY032626	+	44.00	115.75	359.27	15432   AY032626 Porcine reproduct
gb_vi:AF176348	+	44.00	115.72	360.62	15483   AF176348 Porcine reproduct
gb_vi:AF331831	+	44.00	115.71	361.17	15504   AF331831 Porcine reproduct
gb_vi:AF184212	+	44.00	115.70	361.60	15520   AF184212 Porcine reproduct
gb_ba:AF015775	-	44.00	109.88	762.44	29950   AF015775 Bacillus subtilis
gb_ba:AF006665	+	44.00	109.84	766.77	30100   AF006665 Bacillus subtilis
gb_htg:AC091603	+	44.00	102.61	1.9e+03	68129   AC091603 Homo sapiens chr
gb_htg:AC013647	+	44.00	100.99	2.4e+03	81878   AC013647 Homo sapiens clon
gb_htg:LMFLCHX36_27	-	44.00	98.37	3.3e+03	110000   Continuation (28 of 36) o
gb_htg:AC025426	+	44.00	95.42	4.9e+03	153646   AC025426 Homo sapiens chr
gb_htg:CNS07EFP	+	44.00	95.38	4.9e+03	154271   AL513405 Oryza sativa chr
gb_htg:AC079133	+	44.00	95.30	4.9e+03	155622   AC079133 Homo sapiens chr
gb_htg:AL259606	+	44.00	94.64	5.4e+03	176445   AL259606 Homo sapiens chr
gb_pr:CNS01D04	+	44.00	93.83	6.0e+03	183484   AL133163 Human chromosome
gb_htg:AC092482	+	44.00	93.55	6.2e+03	189808   AC092482 Mus musculus clo
gb_htg:AC092261	+	44.00	93.07	6.2e+03	190520   AC092261 Mus musculus clo
gb_htg:AC087442	+	44.00	93.02	6.6e+03	200257   AC087442 Homo sapiens chr
gb_htg:AC073687	+	44.00	92.75	6.9e+03	207730   Z99114 Bacillus subtilis
gb_ba:BSU00011	-	44.00	91.52	8.0e+03	238741   AC088894 Homo sapiens chr
gb_htg:AC068894	+	44.00	90.80	8.8e+03	258629   AC073771 Mus musculus clo
gb_htg:AC073771	+	43.50	92.85	8.8e+03	166464   AL158212 Homo sapiens chr
gb_htg:AL158212	+	43.50	90.85	8.8e+03	208632   AC073711 Mus musculus clo
gb_pat:AX005885	+	43.00	157.52	0.4663	29   AX005885 Sequence 94 from Pat
gb_ba:U32811	+	43.00	117.88	355.98	10064   U32811 Haemophilus influ
gb_ro:MM048854	+	43.00	111.78	597.69	15899   U48854 Mus musculus dystro
gb_pr:AC004051	+	43.00	92.06	7.5e+03	147665   AC004051 Homo sapiens chr
gb_htg:AC026144	+	43.00	92.04	7.5e+03	147990   AC026144 Homo sapiens chr
gb_htg:AC079389	+	43.00	91.82	7.7e+03	151726   AC079389 Rattus norvegicu
gb_pr:AL157384	+	43.00	91.45	8.1e+03	158143   AL157384 Human DNA sequen
gb_htg:AC078940	+	43.00	90.78	8.8e+03	169277   AC078940 Homo sapiens chr
gb_ro:AC074041	+	43.00	90.78	8.8e+03	170548   AC074041 Mus musculus chr
gb_htg:AC016667	+	43.00	90.25	9.5e+03	181079   AC016667 Homo sapiens clo
gb_htg:AL162400	+	43.00	90.03	9.7e+03	185705   AL162400 Homo sapiens chr
gb_htg:AC027180	+	43.00	89.93	9.9e+03	187778   AC027180 Homo sapiens clo
gb_htg:AC03658	+	43.00	89.88	9.9e+03	188873   AC03658 Mus musculus clo
gb_htg:AC025784	+	43.00	89.64	1.0e+04	193661   AC025784 Mus musculus clo
gb_htg:AC091474	+	43.00	89.43	1.1e+04	198653   AC091474 Mus musculus chr
gb_htg:AC091422	+	43.00	88.76	1.1e+04	214440   AC091422 Mus musculus chr
gb_htg:AC093359	+	43.00	88.59	1.2e+04	218502   AC093359 Mus musculus clo
gb_htg:AC092033	+	43.00	88.57	1.2e+04	218946   AC092033 Homo sapiens chr
gb_htg:AC092481	+	43.00	88.18	1.2e+04	228959   AC092481 Mus musculus clo
gb_htg:AC079554	+	43.00	87.70	1.3e+04	241714   AC079554 Mesorhizobium lo
gb_ba:SM053002	-	43.00	84.43	2.0e+04	349498   AP003002 Mesorhizobium me
gb_ba:SM0591792	-	42.50	82.99	2.4e+04	333800   AL591792 Sinorhizobium me
gb_pr:HS26633F	-	42.00	143.19	10.64	300   Z60511 H.sapiens Cpg island
gb_pl:AF361627	-	42.00	133.16	38.50	932   AF361627 Arabidopsis thalian
gb_ro:573473	-	42.00	129.67	60.26	1383   S73473 beta 3-adrenergic re
gb_ro:556481	-	42.00	126.55	89.92	1968   S56481 beta 3-adrenergic re
gb_ro:R56ADR8	-	42.00	125.64	101.04	2181   M74716 Rat beta-adrenergic
gb_in:DROPLC21A	-	42.00	117.37	291.72	5552   M60452 D.melanogaster phosph
gb_in:DROPLC21B	-	42.00	117.06	303.55	5750   M60453 D.melanogaster phosph

gb_ba:AE001520	10516	111.72	602.21	42.00	10516	AE001520 Helicobacter pylori,	gb_ba:AF064682	3664	117.34	292.97	AF064682 Porphyromonas ging
gb_pat:AX191747	10516	111.72	602.21	-	gb_v1:CMU19239	AX191747 Sequence 29 from Pat	gb_v1:CMU19239	4000	117.34	323.64	U19239 Chorisoneura fumifera
gb_ba:AE000606	10719	111.55	615.42	42.00	gb_tg:AE000606	Helicobacter pylori	gb_tg:AE000606	5267	41.00	114.13	U21795 Mus musculus common
gb_pl:AE003904	13254	109.68	783.05	42.00	U03904 Emerichia nidulans cy	U03904 Emerichia nidulans cy	gb_pl:AE003904	8301	41.00	110.10	AF053008 Glycine max env ps
gb_ba:AY005136	14155	109.09	843.73	42.00	AY005136 Chlorobium tepidum F	AY005136 Chlorobium tepidum F	gb_ba:AE007865	8712	41.00	109.68	AF007865 Clostridium acetob
gb_pl:ATU53501	37570	100.46	2.6e+03	42.00	U35501 Arabidopsis thaliana C	U35501 Arabidopsis thaliana C	gb_ba:AE002398	10629	41.00	107.92	AE002398 Nelisseria meningi
gb_ba:AC017345	44868	98.89	3.1e+03	42.00	AC017345 Drosophila melanog	AC017345 Drosophila melanog	gb_ba:AE004239	10704	41.00	107.86	AE004239 Vibrio cholerae c
gb_in:AC004115	80101	92.76	6.0e+03	42.00	AC004115 Drosophila melanog	AC004115 Drosophila melanog	gb_ba:AE005186	12518	41.00	106.47	AE005186 Escherichia coli
gb_pl:AC003040	98374	91.94	7.6e+03	42.00	AC003040 Arabidopsis thaliana	AC003040 Arabidopsis thaliana	gb_vt:LDVGLVPO1	13699	41.00	105.67	AC010529 Homo sapiens chro
gb_pl:F23A5	109694	90.98	8.6e+03	42.00	AC011713 Arabidopsis thaliana	AC011713 Arabidopsis thaliana	gb_vt:LDVGLVPO1	14225	41.00	105.34	L13298 Lactate dehydrogena
gb_pr:AL158152	114517	90.60	9.0e+03	42.00	AL158152 Human DNA sequence	AL158152 Human DNA sequence	gb_ba:AE002550	15053	41.00	104.84	AE002550 Nelisseria meningi
gb_pr:AF000997	130448	89.44	1.0e+04	42.00	AF000997 Homo sapiens DNA, c	AF000997 Homo sapiens DNA, c	gb_ba:AE000736	15862	41.00	104.38	AE000736 Aquifex aeolicus
gb_ba:AC078864	138851	88.89	1.1e+04	42.00	AC078864 Homo sapiens chromo	AC078864 Homo sapiens chromo	gb_ba:AE004132	16251	41.00	104.16	AE004132 Vibrio cholerae c
gb_pr:AC007314	139049	88.46	1.1e+04	42.00	AC007314 Homo sapiens BAC cl	AC007314 Homo sapiens BAC cl	gb_ba:AE000118	21757	41.00	101.58	AE000118 Escherichia coli
gb_hg:AC019244	145770	88.46	1.1e+04	42.00	AC019244 Homo sapiens clone	AC019244 Homo sapiens clone	gb_ba:EC2MIN	28277	41.00	99.26	XS05034 E. coli 2 minute re
gb_hg:CN501DVT	151927	88.09	1.2e+04	42.00	AL136038 Homo sapiens chromo	AL136038 Homo sapiens chromo	gb_pat:AX191720	28277	41.00	99.26	AX191720 Sequence 2 from P
gb_hg:AC036818	153101	88.03	1.3e+04	42.00	AC036818 Homo sapiens chromo	AC036818 Homo sapiens chromo	gb_pr:AC002306	38651	41.00	96.50	AC02306 Homo sapiens DNA
gb_pr:AL450104	154879	87.92	1.4e+04	42.00	AL450104 Human DNA sequence	AL450104 Human DNA sequence	gb_pl:SPAC11D3	39292	41.00	96.35	Z68166 S.pombe chromosome
gb_hg:AC083821	157242	87.79	1.3e+04	42.00	AC083821 Homo sapiens chromo	AC083821 Homo sapiens chromo	gb_hg:AC039084_3	39292	41.00	96.35	Continuation (4 of 4) of A
gb_hg:AC019706	158293	87.73	1.3e+04	42.00	AC019706 Drosophila melanog	AC019706 Drosophila melanog	gb_hg:HSJ393D12	58387	41.00	92.85	AL132776 Human DNA sequen
gb_hg:AC007924	159593	87.66	1.3e+04	42.00	AC007924 Homo sapiens chromo	AC007924 Homo sapiens chromo	gb_pr:AC007445	97137	41.00	88.34	AC007445 Homo sapiens BAC
gb_hg:AL445653	160780	87.59	1.3e+04	42.00	AL445653 Homo sapiens chromo	AL445653 Homo sapiens chromo	gb_pl:SCU12980	103682	41.00	87.76	U12980 Saccharomyces cere
gb_in:AC008371	163461	87.45	1.4e+04	42.00	AC008371 Drosophila melanog	AC008371 Drosophila melanog	gb_hg:AC011458	105412	41.00	87.62	AC011458 Homo sapiens chr
gb_hg:AP003372	164349	87.40	1.4e+04	42.00	AP003372 Oryza sativa chromo	AP003372 Oryza sativa chromo	gb_hg:AC035148	107850	41.00	87.42	AC035148 Homo sapiens chr
gb_pr:CN501DVT	164628	87.38	1.4e+04	42.00	AL135858 Human chromosome 14	AL135858 Human chromosome 14	gb_ba:EC0110K	111401	41.00	87.13	DI0483 E.coli K12 genome, c
gb_hg:AC084819	165452	87.34	1.4e+04	42.00	AC084819 Homo sapiens chromo	AC084819 Homo sapiens chromo	gb_hg:AC010107	112958	41.00	87.01	AC010107 Drosophila melan
gb_hg:AC087784	167310	87.24	1.4e+04	42.00	AC087784 Homo sapiens chromo	AC087784 Homo sapiens chromo	AC027441	116230	41.00	86.75	AC027441 Homo sapiens chr
gb_hg:AP002991	167994	87.21	1.4e+04	42.00	AP002991 Homo sapiens chromo	AP002991 Homo sapiens chromo	AC027441	116230	41.00	86.75	AC027441 Homo sapiens chr
gb_hg:AL356126	170586	87.07	1.4e+04	42.00	AL356126 Homo sapiens chromo	AL356126 Homo sapiens chromo	AP000887	120664	41.00	86.42	AP000887 Homo sapiens gen
gb_hg:AC027624	172118	86.99	1.4e+04	42.00	AC027624 Homo sapiens chromo	AC027624 Homo sapiens chromo	AC079626	122728	41.00	86.27	AC079626 Mus musculus chr
gb_hg:AL351244	172510	86.97	1.4e+04	42.00	AL351244 Homo sapiens chromo	AL351244 Homo sapiens chromo	AL139090	128468	41.00	85.87	AL139090 Human DNA sequen
gb_hg:AC087241	174006	86.89	1.5e+04	42.00	AC087241 Homo sapiens chromo	AC087241 Homo sapiens chromo	AL034343	130904	41.00	85.70	AL034343 Human DNA sequen
gb_hg:AC025622	175093	86.84	1.5e+04	42.00	AC025622 Mus musculus clone	AC025622 Mus musculus clone	AL359096	131299	41.00	85.68	AL359096 Homo sapiens chr
gb_hg:AC035349	175572	86.82	1.5e+04	42.00	AC035349 Homo sapiens chromo	AC035349 Homo sapiens chromo	AC029296	136883	41.00	85.31	AC029296 Homo sapiens chr
gb_hg:AP003902	176877	86.75	1.5e+04	42.00	AP003902 Oryza sativa chromo	AP003902 Oryza sativa chromo	AP000405	141675	41.00	85.00	AP000405 Homo sapiens chr
gb_hg:AC036286	181902	86.50	1.5e+04	42.00	AC036286 Homo sapiens chromo	AC036286 Homo sapiens chromo	AP001038	141895	41.00	84.99	AP001038 Homo sapiens gen
gb_hg:AC034251	184056	86.40	1.6e+04	42.00	AC034251 Homo sapiens chromo	AC034251 Homo sapiens chromo	AC026442	142592	41.00	84.95	AC026442 Homo sapiens chr
gb_in:AC009348	187151	86.25	1.6e+04	42.00	AC009348 Drosophila melanog	AC009348 Drosophila melanog	AC073826	146190	41.00	84.73	AC073826 Mus musculus clo
gb_hg:AL591376	202182	85.57	1.7e+04	42.00	AL591376 Mus musculus clone	AL591376 Mus musculus clone	AC009943	150236	41.00	84.48	AC009943 Homo sapiens clo
gb_hg:AC079529	209301	85.26	1.8e+04	42.00	AC079529 Mus musculus clone	AC079529 Mus musculus clone	AL450103	153914	41.00	84.27	AL450103 Homo sapiens chr
gb_pr:AC028259	210158	85.22	1.8e+04	42.00	AC028259 Homo sapiens 12 BAC	AC028259 Homo sapiens 12 BAC	AC073295	154256	41.00	84.25	AC073295 Mus musculus clo
gb_hg:AC079426	212668	85.12	1.8e+04	42.00	AC079426 Mus musculus chromo	AC079426 Mus musculus chromo	AL359714	156129	41.00	84.14	AL359714 Human DNA sequen
gb_hg:AC073717	215105	85.02	1.9e+04	42.00	AC073717 Mus musculus clone	AC073717 Mus musculus clone	AP002542	156266	41.00	84.14	AP002542 Oryza sativa gen
gb_hg:AC093339	215352	85.01	1.9e+04	42.00	AC093339 Mus musculus clone	AC093339 Mus musculus clone	AC022795	157888	41.00	84.04	AC022795 Homo sapiens chr
gb_hg:AL513345	217178	84.93	1.9e+04	42.00	AL513345 Mus musculus chromo	AL513345 Mus musculus chromo	AC080158	160203	41.00	83.92	AC080158 Mus musculus chr
gb_in:AE000732	234369	84.26	2.0e+04	42.00	AE000732 Drosophila melanog	AE000732 Drosophila melanog	AP000888	166355	41.00	83.58	AP000888 Homo sapiens gen
gb_hg:AL596136	254871	83.52	2.2e+04	42.00	AL596136 Mus musculus chromo	AL596136 Mus musculus chromo	AC06122	166953	41.00	83.55	AC06122 Homo sapiens chr
gb_hg:AL359385	259198	83.37	2.3e+04	42.00	AL359385 Homo sapiens chromo	AL359385 Homo sapiens chromo	AC021774	167366	41.00	83.53	AC021774 Homo sapiens clo
gb_in:AE003589	302473	82.00	2.7e+04	42.00	AE003589 Drosophila melanog	AE003589 Drosophila melanog	AC084290	168249	41.00	83.48	AC084290 Homo sapiens chr
gb_ba:TACID1	344000	80.86	3.1e+04	42.00	AL445063 Thermoplasma acidop	AL445063 Thermoplasma acidop	gb_hg:AL356260	168464	41.00	83.47	AL356260 Homo sapiens chr
gb_pat:116698	235	116.69	12.98	41.00	Sequence 43 from patent	Sequence 43 from patent	gb_hg:AL596444	170658	41.00	83.36	AL596444 Mus musculus chr
gb_pat:116699	235	116.69	12.98	41.00	Sequence 45 from patent	Sequence 45 from patent	gb_hg:AC024662	171549	41.00	83.31	AC024662 Homo sapiens chr
gb_pat:AX005865	285	116.69	12.98	41.00	Sequence 74 from Patent	Sequence 74 from Patent	gb_pr:AC015973	171834	41.00	83.30	AC015973 Homo sapiens BAC
gb_hg:AC089106	310	139.94	16.15	41.00	AX089106 Giardia intestinalis c	AX089106 Giardia intestinalis c	gb_pat:AX044035	172325	41.00	83.27	AX044035 Sequence 114 fro
gb_r:S75844S8	468	135.55	28.36	41.00	S75852 IL-2R gamma -interleukin	S75852 IL-2R gamma -interleukin	gb_hg:AC073748	175371	41.00	83.11	AC073748 Mus musculus clo
gb_ba:D78201	716	131.79	45.94	41.00	D78201 Borrelia miyamotoi gene	D78201 Borrelia miyamotoi gene	gb_hg:AC087868	178676	41.00	82.95	AC087868 Mus musculus clo
gb_ba:AE001719	717	131.77	46.02	41.00	Borrelia miyamotoi gene	Borrelia miyamotoi gene	gb_hg:AC024000	179032	41.00	82.93	AC024000 Homo sapiens chr
gb_sts:CN506KJX	978	129.03	65.45	41.00	AB02979 T7 end of clone AT0A0	AB02979 T7 end of clone AT0A0	gb_hg:AC032809	179376	41.00	82.92	AC032809 Homo sapiens chr
gb_ba:S81109	1301	126.50	90.48	41.00	S81109 citrate synthase [Pyroc	S81109 citrate synthase [Pyroc	gb_hg:AC007961	179637	41.00	82.90	AC007961 Mus musculus chr
gb_ba:AF015823	1404	125.83	98.65	41.00	AX015823 Streptomyces venezue	AX015823 Streptomyces venezue	gb_hg:AP001352	181413	41.00	82.82	AP001352 Homo sapiens chr
gb_pat:AX189061	1476	125.39	104.41	41.00	AX189061 Sequence 262 from Pat	AX189061 Sequence 262 from Pat	gb_hg:AL360090	181936	41.00	82.79	AL360090 Homo sapiens chr
gb_r:MUSIL2RGA	1590	124.73	113.61	41.00	L13565 Mus musculus mRNA for i	L13565 Mus musculus mRNA for i	gb_hg:AC092215	183533	41.00	82.71	AC092215 Mus musculus clo
gb_pat:AR072026	1608	124.63	115.07	41.00	AR072026 Sequence 1 from patent	AR072026 Sequence 1 from patent	gb_hg:AC084402	184747	41.00	82.65	AC084402 Mus musculus chr
gb_r:MUSIL2RG	1608	124.63	115.07	41.00	Mouse Interleukin 2 rec	Mouse Interleukin 2 rec	AP000063	185300	41.00	82.63	AP000063 Aeropyrum pernix
gb_r:MMML2RG	1623	124.55	116.29	41.00	X75337 M.musculus mRNA for mur	X75337 M.musculus mRNA for mur	AC073561	185699	41.00	82.61	AC073561 Mus musculus chr
gb_ba:ECMURG	1713	124.07	123.63	41.00	X52540 Escherichia coli cell-e	X52540 Escherichia coli cell-e	gb_hg:AC079557	185938	41.00	82.60	AC079557 Homo sapiens clo
gb_r:ECMURG	1776	123.75	128.80	41.00	U95086 Mus musculus AtptA gene	U95086 Mus musculus AtptA gene	AP000487	190218	41.00	82.40	AP000487 Homo sapiens chr
gb_ba:ECU67892	1998	122.71	147.22	41.00	U67892 Escherichia coli UDP-Mu	U67892 Escherichia coli UDP-Mu	gb_hg:AC092628	190756	41.00	82.37	AC092628 Homo sapiens chr
gb_pat:AX166563	2730	119.94	209.80	41.00	AX166563 Sequence 54 from Pat	AX166563 Sequence 54 from Pat	gb_hg:AL365326	195603	41.00	82.15	AL365326 Mus musculus chr
gb_ba:ECMURG	2793	119.74	215.30	41.00	X52644 Escherichia coli murG a	X52644 Escherichia coli murG a	gb_ba:NMA72491	195767	41.00	82.14	AL162758 Nelisseria mening
gb_pr:AC013118	2839	119.60	219.33	41.00	AB033118 Homo sapiens mRNA for	AB033118 Homo sapiens mRNA for	gb_pr:AC013118	196721	41.00	82.10	AC013118 Homo sapiens BAC
gb_pl:YSCFN12A	3075	118.89	240.14	41.00	L29389 Saccharomyces cerevisia	L29389 Saccharomyces cerevisia	gb_hg:AC013946	197232	41.00	82.08	AC013946 Mus musculus chr

gb_hhg:AL355513	197452	1	AL355513	Homo sapiens chromo	40.00	103.50	1.7e+03	11513	gb_to:AF353728	Mus musculus myo-
gb_hhg:AL591711	200875	1	AL591711	Mus musculus chromo	40.00	101.50	2.0e+03	12879	gb_ba:AE004609	Pseudomonas aerugin
gb_pr:AL351833	201512	1	AL351833	Human DNA sequenc	40.00	102.71	2.2e+03	14104	gb_vi:LDU15146	Lactate dehydrogena
gb_hlg:AC084827	202052	1	AC084827	Mus musculus clone	40.00	98.96	3.1e+03	19227	gb_pat:AR18085	Sequence 13 from
gb_hlg:AC010413	203876	1	AC010413	Homo sapiens chromo	40.00	98.59	3.2e+03	20062	gb_ba:AE003919	Xylella fastidios
gb_hlg:AC073749	206533	1	AC073749	Mus musculus clone	40.00	97.83	3.6e+03	21852	gb_pl:SPAC3D6	S.pombe chromosome
gb_hlg:AC079988	208378	1	AC079988	Homo sapiens chromo	40.00	96.80	4.0e+03	24000	gb_ba:AE004856	Pseudomonas aerugin
gb_hlg:AC073828	213574	1	AC073828	Mus musculus clone	40.00	97.00	4.1e+03	24550	gb_in:CEF13G3	271259 Caenorhabditis eleg
gb_hlg:AC087153	216133	1	AC087153	Mus musculus clone	40.00	96.73	4.1e+03	24753	gb_pr:HS329F2	AL031710 Human DNA sequenc
gb_hlg:AC079272	227366	1	AC079272	Mus musculus chromo	40.00	95.61	4.8e+03	28080	gb_in:AF016438	Caenorhabditis el
gb_hlg:AL591514	230933	1	AL591514	Mus musculus chromo	40.00	95.02	5.1e+03	30027	gb_in:CELT2AD11	U49940 Caenorhabditis eleg
gb_hlg:AL596103	232712	1	AL596103	Mus musculus chromo	40.00	92.73	6.9e+03	38875	gb_pr:AC006950	Homo sapiens chro
gb_ba:AP000003	233000	1	AP000003	Pyrococcus horikosh	40.00	92.45	7.1e+03	40138	gb_pat:AR118084	Sequence 12 from
gb_hlg:AC073760	237653	1	AC073760	Mus musculus clone	40.00	92.14	7.4e+03	41587	gb_ba:AB025342	AB025342 Moritella marina
gb_hlg:AC093263	241309	1	AC093263	Homo sapiens chromo	40.00	91.63	7.9e+03	44042	gb_pl:SPCC736	AL023705 S.pombe chromosom
gb_ro:AC079869	243675	1	AC079869	Mus musculus 9 BAC	40.00	87.72	1.3e+04	68484	gb_hlg:AL162408	AL592437 Homo sapiens chr
gb_hlg:AC090887	24849	1	AC090887	Mus musculus clone	40.00	86.24	1.6e+04	80993	gb_in:AC004313	AL355926 Neurospora crassa
gb_hlg:AC073822	247196	1	AC073822	Mus musculus clone	40.00	85.96	1.6e+04	83551	gb_hlg:AC018150	AC004313 Drosophila melano
gb_ba:AP000250	281530	1	AP000250	Escherichia coli O1	40.00	85.70	1.7e+04	86043	gb_hhg:AC084694	AC018150 Drosophila melano
gb_ba:CNBPAX05	304290	1	AJ248287	Pyrococcus abyssi c	40.00	85.15	1.8e+04	91570	gb_hlg:ATTT13K14	AL080282 Arabidopsis thali
gb_hlg:AC093082	326924	1	AC093082	Homo sapiens chromo	40.00	84.33	2.0e+04	100499	gb_pr:AR138769	AL138769 Human DNA sequen
gb_pr:AP001732	340000	1	AP001732	Homo sapiens genom	40.00	84.17	2.1e+04	102308	gb_pr:AC005231	AC005231 Homo sapiens PAC
gb_ba:AP003001	345783	1	AP003001	Mesorhizobium loti	40.00	83.53	2.2e+04	110000	gb_hhg:AC073702_2	Continuation (3 of 5) of
gb_ba:AP001118	347550	1	AP001118	Buchnera sp. APS ge	40.00	83.07	2.4e+04	115932	gb_pr:HS107	AL031589 Human DNA sequen
gb_ba:NMA222491	349061	1	AL162753	Neisseria meningit	40.00	82.91	2.4e+04	117955	gb_pr:AC019052	AC019052 Homo sapiens clo
gb_pat:AX041921	349980	1	AX041921	Sequence 816 from F	40.00	82.84	2.4e+04	118929	gb_hhg:AC025976	AC025976 Homo sapiens chr
gb_pat:AX044029	349980	1	AX044029	Sequence 108 from F	40.00	82.84	2.4e+04	118929	gb_hhg:AC025976	AC025976 Homo sapiens chr
gb_pat:AX044034	349980	1	AX044034	Sequence 113 from F	40.00	82.73	2.5e+04	120423	gb_hlg:HS3261L2	AL023279 Homo sapiens DNA
gb_in:AF189278	1611	1	AF189278	Drosophila melanogast	40.00	82.44	2.6e+04	124497	gb_pr:HS232G24	AL022152 Homo sapiens clo
gb_ba:AE008219	9588	1	AE008219	Agrobacterium tumefac	40.00	82.19	2.7e+04	127934	AC073709	Mus musculus clo
gb_hlg:AC007995	204206	1	AC007995	Mus musculus clone	40.00	81.77	2.8e+04	134222	AB045363	Homo sapiens gen
gb_in:AE003516	287649	1	AE003516	Drosophila melano	40.00	81.42	2.9e+04	136665	AP000865	Homo sapiens chr
gb_pl:VNA401056	222	1	AJ401056	Vicia narbonensis Tyl-	40.00	80.54	3.3e+04	154309	AL592437	Homo sapiens chr
gb_pat:AX005861	230	1	AX005861	Sequence 70 from Pat	40.00	80.26	3.4e+04	159173	AC015992	Homo sapiens chr
gb_sts:AU027871	251	1	AU027871	Rattus norvegicus, OR	40.00	80.21	3.4e+04	160125	AC016947	Homo sapiens chr
gb_ro:RAT5A117	395	1	L00124	Rat elastase II gene, ex	40.00	80.15	3.4e+04	161175	AC073758	Mus musculus clo
gb_sts:G25729	423	1	G25729	human STS EST156852, seq	40.00	80.09	3.5e+04	162269	AC000097	Homo sapiens chr
gb_hlg:AC037213	453	1	G29444	human STS SHGC-33656, se	40.00	80.02	3.5e+04	163559	AC018635	Homo sapiens chr
gb_hlg:AC037213	603	1	AC037213	Giardia intestinalis c	40.00	79.94	3.5e+04	165073	AC073516	Homo sapiens chr
gb_pat:AX142463	603	1	AX142463	Sequence 1185 from Pat	40.00	79.93	3.5e+04	165243	AL162408	Homo sapiens chr
gb_hlg:AC028609	819	1	AX143667	Sequence 2389 from Pat	40.00	79.86	3.6e+04	166564	AC026101	Homo sapiens chr
gb_hlg:AC028609	819	1	AC028609	Giardia intestinalis c	40.00	79.85	3.6e+04	166751	AC087066	Rattus norvegicu
gb_hlg:AC061108	918	1	AC061108	Giardia intestinalis c	40.00	79.84	3.6e+04	166892	AC007463	Homo sapiens BAC
gb_hlg:AC037213	954	1	AC037213	Giardia intestinalis c	40.00	79.80	3.6e+04	167691	AL023284	Human DNA sequen
gb_hlg:AC031132	982	1	AC031132	Giardia intestinalis	40.00	79.79	3.6e+04	167739	AC020547	Homo sapiens cto
gb_hlg:AC080985	1059	1	AC080985	Giardia intestinalis	40.00	79.73	3.6e+04	168990	AC064834	Homo sapiens chr
gb_hlg:AC047651	1086	1	AC047651	Giardia intestinalis	40.00	79.63	3.7e+04	170983	AL590117	Homo sapiens chr
gb_hlg:AC047850	1090	1	AC047850	Giardia intestinalis	40.00	79.57	3.7e+04	172114	AP001999	Homo sapiens chr
gb_hlg:AC047850	1101	1	AC047850	Giardia intestinalis	40.00	79.54	3.7e+04	172797	AC011298	Homo sapiens chr
gb_ba:SS0012093	1424	1	AC047849	Sulfolobus solfataric	40.00	79.43	3.8e+04	174840	AC006549	Homo sapiens chr
gb_ro:BC003930	1424	1	AC07849	Giardia intestinalis	40.00	79.38	3.8e+04	175925	AC068110	Homo sapiens chr
gb_hlg:AC06766	1454	1	BC003930	Mus musculus, RIKEN C	40.00	79.23	3.9e+04	178944	AP000804	Homo sapiens chr
gb_ba:AF110740	1700	1	AC06766	Pseudomonas aeruginos	40.00	79.21	3.9e+04	179269	AC006547	Homo sapiens chr
gb_pr:575415	2191	1	AF110740	Pseudomonas aeruginos	40.00	79.14	3.9e+04	180742	AL359753	Human DNA sequen
gb_com:BOBACAM	2593	1	S75415	IF16-interferon-induct	40.00	79.12	3.9e+04	181047	AC016765	Homo sapiens chr
gb_pl:AF27068	3035	1	X12672	Bovine mRNA for oploid	40.00	79.08	3.9e+04	181902	AC020788	Homo sapiens chr
gb_ba:AF270288	3246	1	AF27068	Arabidopsis thaliana	40.00	79.38	4.0e+04	184180	AL589623	Mus musculus chr
gb_pat:AX145606	3246	1	AF270288	Staphylococcus epide	40.00	78.97	4.0e+04	184433	AL596446	Mus musculus chr
gb_ba:AF269682	3645	1	AX145606	Sequence 4328 from Pa	40.00	78.95	4.0e+04	184651	AL450317	Mus musculus chr
gb_pat:AX145002	3645	1	AF269682	Staphylococcus epide	40.00	78.93	4.0e+04	184929	AC058790	Homo sapiens chr
gb_ba:AF269845	4248	1	AX145002	Sequence 3724 from Pa	40.00	78.85	4.1e+04	186729	AC01296	Pan troglodytes
gb_pat:AX145163	557	1	AF269845	Staphylococcus epide	40.00	78.79	4.1e+04	188034	AC036169	Homo sapiens chr
gb_ov:SANKCC1	5260	1	AX145163	Sequence 3885 from Pa	40.00	78.76	4.1e+04	188561	AL592488	Homo sapiens chr
gb_vi:AF201894	7374	1	U05958	Squalus acanthias bumet	40.00	78.71	4.1e+04	189612	AC051617	Mus musculus cto
gb_pl:AF060248	7894	1	AF201894	A-2 plaque virus A2 H	40.00	78.57	4.2e+04	192782	AC079276	Mus musculus chr
gb_ba:AE005138	10329	1	AF060248	Arabidopsis thaliana	40.00	78.44	4.3e+04	195630	AC087898	Mus musculus chr
gb_ba:AE006777	10640	1	AE005138	Halobacterium sp. NH	40.00	78.42	4.3e+04	195931	AC068174	Homo sapiens chr
gb_ba:AE007899	10651	1	AE006777	Sulfolobus solfatar	40.00	78.39	4.3e+04	196606	AP000779	Homo sapiens gen
gb_ba:AE004474	10994	1	AE007899	Agrobacterium tumefa	40.00	78.39	4.3e+04	196722	AC092481	Mus musculus cto
gb_ba:AE007244	10993	1	AE004474	Pseudomonas aerugin	40.00	78.35	4.3e+04	197568	AL161553	Arabidopsis thal
gb_ba:AE006890	11154	1	AE007244	Sinorhizobium melioid	40.00	78.30	4.3e+04	198602	AL450395	Mouse DNA sequen
gb_ba:HHRNAPO	11208	1	AE006890	Sulfolobus solfatar	40.00	78.30	4.3e+04	198739	AC079840	Homo sapiens chr
		1	X57144	H. halobium RNA polym	40.00	78.30	4.3e+04	198739	AC079840	Homo sapiens chr

gb_htg:AC074027	199359	! AC074027	Mus musculus chromo	gb_htg:AC018135	-	39.00	108.58	901.38	4264	! AC018135	Drosophila melanog
gb_htg:AC034116	200254	! AC034116	Mus musculus clone	gb_ov:DFGAGR	-	39.00	108.39	923.00	4354	! M04123	Discothyris ommata (cl
gb_htg:AL596386	200720	! AL596386	Mus musculus chromo	gb_ov:EGCOPATBCD	-	39.00	108.33	930.46	4385	! M45159	E.coli transport pro
gb_htg:AC064868	200859	! AC064868	Homo sapiens chromo	gb_ba:AF077856	+	39.00	108.31	932.87	4395	! AF077856	Actinobacillus act
gb_pr:AP001256	204699	! AP001256	Homo sapiens genom	gb_ba:AF077856	+	39.00	108.31	932.87	4395	! AF077856	Actinobacillus act
gb_htg:AC073792	204929	! AC073792	Mus musculus clone	gb_pr:AK023832	-	39.00	108.17	950.23	4467	! E33370	NA+ATPase gene. 2/2
gb_htg:AL591032	205676	! AL591032	Mus musculus chromo	gb_pr:CRUKLGG1	-	39.00	107.91	981.43	4596	! AC073832	Homo sapiens cDNA
gb_htg:AL591964	206191	! AL591964	Mus musculus chromo	gb_pr:ATU63931	+	39.00	107.86	987.98	4623	! L47236	Cricetulus griseus (
gb_htg:AC074161	210539	! AC074161	Mus musculus clone	gb_pl:ATU63931	+	39.00	107.79	997.44	4662	! U63931	Arabidopsis thaliana
gb_htg:AC025751	214894	! AC025751	Mus musculus clone	gb_ba:ASU21850	+	39.00	107.71	1.0e+03	4704	! AB028990	Homo sapiens mRNA
gb_htg:AC073812	219043	! AC073812	Mus musculus clone	gb_ba:ASU21850	+	39.00	107.28	1.1e+03	4938	! U21853	Anabaena sp. phycob
gb_htg:AC063968	220237	! AC063968	Mus musculus clone	gb_pr:AB078736	+	39.00	106.36	1.2e+03	5480	! AB078736	Homo sapiens mRNA
gb_htg:AC019008	224956	! AC019008	Mus musculus chromo	gb_pr:AB078736	+	39.00	105.66	1.3e+03	5932	! AB07916	Homo sapiens mRNA
gb_htg:AC073704	226153	! AC073704	Mus musculus clone	gb_ba:PPORVOR	+	39.00	101.41	2.3e+03	5985	! X85250	P.furiosus vorA, vor
gb_htg:AC060772	227074	! AC060772	Mus musculus clone	gb_ba:AE004815	+	39.00	101.02	2.4e+03	10018	! AE001782	Thermotoga mariti
gb_htg:AC068561	230899	! AC068561	Mus musculus chromo	gb_ba:AE004815	+	39.00	100.81	2.4e+03	10254	! AE004815	Pseudomonas aerug
gb_htg:AC073770	231661	! AC073770	Mus musculus clone	gb_ba:AE005326	+	39.00	100.70	2.5e+03	10382	! AE005326	Escherichia coli
gb_htg:AL513352	235039	! AL513352	Mus musculus chromo	gb_ba:AE001865	+	39.00	100.67	2.5e+03	10419	! AE001865	Deinococcus radio
gb_htg:AC073791	239574	! AC073791	Mus musculus clone	gb_ba:AE001417	+	39.00	100.31	2.6e+03	10856	! AE001417	Vibrio cholerae c
gb_htg:AC051620	249433	! AC051620	Mus musculus clone	gb_ba:AE007818	+	39.00	100.23	2.6e+03	10956	! AE007818	Clostridium aceto
gb_htg:AL596255	250327	! AL596255	Mus musculus chromo	gb_ba:AE005011	+	39.00	100.09	2.7e+03	11131	! AE005011	Halobacterium sp.
gb_htg:AC073495	302547	! AC073495	Mus musculus chromo	gb_ba:AE004221	+	39.00	99.88	2.8e+03	11401	! AE004221	Vibrio cholerae c
gb_htg:AC073495	318930	! AC073495	Mus musculus chromo	gb_ba:U32883	+	39.00	99.66	2.8e+03	11681	! U32813	Hemophilus influen
gb_ba:NM022491	326301	! AL162757	Neisseria meningiti	gb_ba:U32883	+	39.00	99.61	2.8e+03	11749	! AE000212	Escherichia coli
gb_pr:AC016678	340000	! AC016678	Homo sapiens genom	gb_ba:D90748	+	39.00	99.48	2.9e+03	11919	! AB009371	Mus musculus Tsc2
gb_htg:AC025110	79685	! AC025110	Homo sapiens chromo	gb_ro:AB009371	+	39.00	99.48	2.9e+03	12997	! AE006061	Pasteurella multo
gb_htg:AC073763.2	110000	! AC073763.2	Continuation (3 of 4) of AC0	gb_ba:AE006061	+	39.00	98.72	3.3e+03	13409	! AE004466	Pseudomonas aerug
gb_htg:LMF1CHR16_04	110000	! Continuation (5 of 11) of LM		gb_ba:AE008126	+	39.00	98.44	3.3e+03	14545	! AE008126	Agrobacterium tum
gb_htg:AC073740	186243	! AC073740	Mus musculus clone	gb_ba:AE008126	+	39.00	97.72	3.6e+03	14838	! AE008126	Agrobacterium tum
gb_htg:AC091362	238638	! AC091362	Rattus norvegicus c	gb_ba:AE001213	+	39.00	97.55	3.7e+03	14838	! AE001213	Treponema pallidu
gb_pl:GHCA454	240	! Z27350	G.hirsutum (AD45A) copia	gb_pr:AC073495	+	39.00	97.55	3.7e+03	14838	! AX191756	Sequence 38 from
gb_pr:GHCA454	240	! Z27350	G.hirsutum (AD45A) copia	gb_ba:D90748	+	39.00	97.44	3.8e+03	15007	! D90748	Escherichia coli ge
gb_pr:W76274	298	! W76274	Sequence 7 from patent	gb_in:AE003364	+	39.00	97.44	3.8e+03	15042	! AC003364	Drosophila melano
gb_om:SSU55958	303	! U55858	Sus scrofa beta 3 adrene	gb_ov:GGA205458	+	39.00	96.86	4.1e+03	16029	! AJ250458	Gallus gallus zw1
gb_pr:SF29249	375	! S79249	Ig-beta/B29-CD79b (alter	gb_in:CEV10666G	+	39.00	94.30	5.3e+03	20244	! AL10470	Streptomyces coel
gb_pr:HSB291GBV	378	! X83539	H.sapiens B29 mRNA for I	gb_ba:SCF85	+	39.00	94.30	5.3e+03	21420	! X72850	Sphingomyces sp. dx
gb_pr:AC018071	548	! AC018071	Sequence 19 from Patent	gb_ba:SF85	+	39.00	93.97	5.9e+03	22233	! X72850	Sphingomyces sp. dx
gb_ba:AVIFDXA	603	! J03521	A.vinelandii ferredoxin	gb_in:LMFL5856	+	39.00	93.44	6.3e+03	23600	! AF17592	Leishmania major
gb_htg:AC129173	748	! AR129173	Sequence 31 from patent	gb_ba:LMFL5856	+	39.00	89.11	1.1e+04	38510	! AF179611	Zymomonas mobilis
gb_htg:AC051996	881	! AC051996	Giardia intestinalis c	gb_in:CEC0999	+	39.00	88.19	1.2e+04	42703	! Z81465	Caenorhabditis eleg
gb_htg:AC071566	998	! AC071566	Giardia intestinalis c	gb_pr:AC036219	+	39.00	85.37	1.8e+04	49999	! AX015913	Caenorhabditis eleg
gb_htg:AC033878	1000	! AC033878	Giardia intestinalis c	gb_htg:AC036219	+	39.00	85.37	1.8e+04	58779	! AC036219	Homo sapiens chro
gb_ro:MM0309317	1008	! AJ309317	Mus musculus partial	gb_pr:AC073996	+	39.00	84.99	1.9e+04	61359	! AC073996	Homo sapiens chro
gb_htg:AC078596	1025	! AC078596	Giardia intestinalis	gb_htg:AC087185	+	39.00	84.99	1.9e+04	63359	! AC087185	Homo sapiens chro
gb_htg:AC071721	1031	! AC071721	Giardia intestinalis	gb_htg:AL5353144	+	39.00	83.72	2.2e+04	70788	! AL535314	Homo sapiens chro
gb_pl:AF155932	1045	! AF155932	Avena sativa 1.3-beta	gb_htg:AC036189	+	39.00	83.72	2.2e+04	71842	! AC036189	Homo sapiens chro
gb_pr:W76270	1164	! W76270	Sequence 3 from patent	gb_pr:AL359692	+	39.00	82.82	2.5e+04	78376	! AL359692	Human DNA sequenc
gb_ro:BC002181	1254	! BC002181	Mus musculus, clone M	gb_htg:AC023239	+	39.00	82.47	2.6e+04	81557	! AC023239	Mus musculus clon
gb_ro:AF193027	1286	! AF193027	Mus musculus beta-3b	gb_pr:AL593397	+	39.00	82.35	2.6e+04	82646	! AL593397	Human DNA sequenc
gb_pr:BC002843	1316	! BC002843	Homo sapiens, clone M	gb_pl:AB013392	+	39.00	82.19	2.7e+04	84129	! AB013392	Arabidopsis thall
gb_pr:W76271	1360	! W76271	Sequence 4 from patent	gb_pr:AC006473	+	39.00	82.03	2.7e+04	85691	! AC006473	Homo sapiens chro
gb_pr:BC010500	1420	! BC010500	Homo sapiens, clone I	gb_pr:AL445470	+	39.00	81.96	2.7e+04	86332	! AL445470	Human DNA sequenc
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AL445470	+	39.00	81.96	2.7e+04	86332	! AL445470	Human DNA sequenc
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AC004482	+	39.00	80.00	3.5e+04	107815	! AC004482	Homo sapiens chro
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AC004482	+	39.00	80.00	3.5e+04	107815	! AC004482	Homo sapiens chro
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AL356434	+	39.00	79.88	3.6e+04	109224	! AL356434	Homo sapiens chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:HS28283	+	39.00	79.77	3.6e+04	110608	! AL31282	Human DNA sequen
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:DMR29G2	+	39.00	79.61	3.7e+04	112702	! AL13487	Drosophila melan
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:CHS059TEL	+	39.00	79.42	3.8e+04	115087	! AL359233	Human chromosome
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AL391803	+	39.00	79.20	3.9e+04	117951	! AL391803	Human DNA sequen
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AP001482	+	39.00	79.18	3.9e+04	118269	! AP001482	Homo sapiens chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AC005095	+	39.00	79.06	3.9e+04	119937	! AC005095	Homo sapiens BAC
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AC009829	+	39.00	78.99	4.0e+04	120829	! AC009829	Homo sapiens chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AC014962	+	39.00	78.83	4.1e+04	123004	! AC014962	Drosophila melan
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AC014962	+	39.00	78.83	4.1e+04	123004	! AC014962	Drosophila melan
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AL445251	+	39.00	78.76	4.1e+04	123482	! AL445251	Human DNA sequen
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AP001259	+	39.00	78.76	4.1e+04	123953	! AP001259	Homo sapiens chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AC008840	+	39.00	78.47	4.2e+04	126746	! AC008840	Homo sapiens chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AC087144	+	39.00	78.47	4.3e+04	128185	! AC087144	Mus musculus clo
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:AL136179	+	39.00	78.43	4.3e+04	128728	! AL136179	Human DNA sequen
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_htg:AC079417	+	39.00	78.31	4.3e+04	130540	! AC079417	Mus musculus chr
gb_pr:AX151232	1652	! AX151232	Sequence 17 from Paten	gb_pr:HS01761K5	+	39.00	78.21	4.4e+04	131974	! AC050350	Human DNA sequen



gb_htg:AC034267	131980	4.4e+04	78.21	39.00	131980	AC034267	Homo sapiens chromd	gb_htg:AC018665	39.00	75.14	6.5e+04	186632	AC018665	Homo sapiens chr
gb_htg:AC090880	133520	4.5e+04	78.11	39.00	133520	AC090880	Mus musculus clone	gb_htg:AC093353	39.00	75.10	6.5e+04	187614	AC093353	Mus musculus cto
gb_htg:AC091502	138221	4.6e+04	77.80	39.00	138221	AC091502	Rattus norvegicus c	gb_pr:AC006369	39.00	75.06	6.6e+04	188359	AC006369	Homo sapiens BAC
gb_htg:AL592525	143068	4.8e+04	77.50	39.00	143068	AL592525	Homo sapiens chromd	gb_htg:AC022546	39.00	75.04	6.6e+04	188781	AC022546	Homo sapiens chr
gb_htg:AC068266	149235	5.0e+04	77.12	39.00	149235	AC068266	Homo sapiens chromd	gb_htg:AC067817	39.00	75.03	6.6e+04	189027	AC067817	Homo sapiens chr
gb_htg:AL160169	149356	5.1e+04	77.11	39.00	149356	AL160169	Homo sapiens chromd	gb_htg:AC091412	39.00	75.03	6.6e+04	189036	AC091412	Homo sapiens chr
gb_htg:AL596257	154145	5.1e+04	76.96	39.00	154145	AL596257	Homo sapiens chromd	gb_htg:AL596285	39.00	75.03	6.6e+04	189135	AL596285	Mus musculus chr
gb_pr:AC068722	154124	5.2e+04	76.99	39.00	154124	AC068722	Homo sapiens chromd	gb_htg:AC011193	39.00	75.02	6.6e+04	189319	AC011193	Homo sapiens chr
gb_htg:AC091620	152906	5.2e+04	76.91	39.00	152906	AC091620	Papio cynocephalus	gb_htg:AC079270	39.00	75.00	6.6e+04	189709	AC079270	Homo sapiens chr
gb_htg:AC011974	153202	5.2e+04	76.89	39.00	153202	AC011974	Homo sapiens clone	gb_pr:AC010745	39.00	74.99	6.6e+04	189894	AC010745	Homo sapiens BAC
gb_htg:AC087262	154441	5.2e+04	76.82	39.00	154441	AC087262	Rattus norvegicus c	gb_htg:AC021177	39.00	74.96	6.6e+04	190488	AC021177	Homo sapiens chr
gb_pr:D90917	154619	5.3e+04	76.81	39.00	154619	D90917	Synechocystis sp. PC	gb_htg:AC021177	39.00	74.89	6.6e+04	192000	AC021177	Homo sapiens chr
gb_htg:AC009271	155157	5.3e+04	76.78	39.00	155157	AC009271	Homo sapiens chromd	em_htg:hum:AC009271	39.00	74.89	6.6e+04	192000	AC009271	Homo sapiens chr
gb_pr:CNS01DWF	155742	5.3e+04	76.74	39.00	155742	AL137190	Human chromosome 14	em_htg:AC024545	39.00	74.83	6.8e+04	193261	AC024545	Homo sapiens chr
gb_htg:AC050516	157577	5.4e+04	76.64	39.00	157577	AL359212	Human chromosome 14	gb_htg:AC013775	39.00	74.81	6.8e+04	193700	AC013775	Mus musculus cto
gb_htg:AP003443	160569	5.5e+04	76.47	39.00	160569	AP003443	Oryza sativa chromd	gb_htg:AC019430	39.00	74.77	6.8e+04	194609	AC019430	Mus musculus chr
gb_htg:AC023586	160569	5.5e+04	76.47	39.00	160569	AC023586	Homo sapiens chromd	gb_htg:AC079419	39.00	74.75	6.8e+04	195107	AL590430	Mus musculus chr
gb_pr:AC090937	160596	5.5e+04	76.47	39.00	160596	AC090937	Homo sapiens chromd	gb_htg:AC079419	39.00	74.71	6.8e+04	196003	AC079419	Mus musculus chr
gb_pr:AC008511	160770	5.5e+04	76.46	39.00	160770	AC008511	Homo sapiens chromd	gb_htg:HS651N20	39.00	74.70	6.9e+04	196247	AL117346	Homo sapiens chr
gb_htg:AC026007	162408	5.6e+04	76.37	39.00	162408	AC026007	Homo sapiens chromd	gb_htg:AL359176	39.00	74.66	6.9e+04	197195	AL359176	Homo sapiens chr
gb_pr:AC016574	163183	5.6e+04	76.33	39.00	163183	AC016574	Homo sapiens chromd	gb_htg:AC010097	39.00	74.61	6.9e+04	198278	AC010097	Homo sapiens chr
gb_htg:AC026761	164513	5.6e+04	76.26	39.00	164513	AC026761	Mus musculus chromd	gb_htg:AL591436	39.00	74.59	7.0e+04	198780	AL591436	Mus musculus chr
gb_pr:HSJ29318	164808	5.6e+04	76.24	39.00	164808	AL078594	Human DNA sequence	gb_htg:AL353729	39.00	74.54	7.0e+04	199762	AL353729	Homo sapiens chr
gb_htg:AL353093	164845	5.6e+04	76.24	39.00	164845	AL353093	Homo sapiens chromd	gb_htg:AC092797	39.00	74.49	7.1e+04	200911	AC092797	Homo sapiens chr
gb_pr:CNS01DMC	166486	5.7e+04	76.15	39.00	166486	AL137100	Human chromosome 14	gb_htg:AC079442	39.00	74.47	7.1e+04	201289	AC079442	Mus musculus chr
gb_htg:AC016165	166578	5.7e+04	76.14	39.00	166578	AC016165	Homo sapiens chromd	gb_htg:AL591706	39.00	74.46	7.1e+04	201725	AL591706	Homo sapiens chr
gb_htg:AC025318	167277	5.7e+04	76.11	39.00	167277	AC025318	Homo sapiens chromd	gb_pr:CNS05TCL	39.00	74.45	7.1e+04	201844	AL355102	Human chromosome
gb_pr:AC079468	167656	5.8e+04	76.09	39.00	167656	AC079468	Homo sapiens chromd	gb_htg:AL596117	39.00	74.45	7.1e+04	201852	AC061963	Mus musculus chr
gb_htg:AC053483	168332	5.8e+04	76.06	39.00	168332	AC053483	Homo sapiens chromd	gb_htg:AC021667	39.00	74.43	7.1e+04	202362	AL596117	Mus musculus chr
gb_htg:AC062007	168466	5.8e+04	76.05	39.00	168466	AC062007	Homo sapiens chromd	gb_ro:AC021667	39.00	74.39	7.1e+04	203106	AC021667	Mus musculus cl
gb_in:AC003463	168493	5.8e+04	76.05	39.00	168493	AC013463	Homo sapiens chromd	gb_htg:AC021445	39.00	74.37	7.2e+04	203621	AC021445	Mus musculus cto
gb_pr:AL359997	169102	5.8e+04	76.02	39.00	169102	AL359997	Human DNA sequence	gb_ro:AC079644	39.00	74.32	7.2e+04	204394	AC079644	Mus musculus 12
gb_htg:AC091503	169514	5.8e+04	75.99	39.00	169514	AC091503	Rattus norvegicus c	gb_htg:AC010359	39.00	74.34	7.2e+04	204943	AC010359	Homo sapiens chr
gb_pr:AC064869	169514	5.8e+04	75.99	39.00	169514	AC064869	Homo sapiens clone	gb_htg:AC087227	39.00	74.28	7.2e+04	205691	AC087227	Mus musculus chr
gb_htg:AC016481	169913	5.8e+04	75.97	39.00	169913	AC016481	Homo sapiens clone	gb_htg:AC004487	39.00	74.27	7.2e+04	205910	AC004487	Mus musculus chr
gb_htg:AC069447	170105	5.9e+04	75.96	39.00	170105	AC069447	Mus musculus chromd	gb_htg:AC087135	39.00	74.10	7.4e+04	209958	AC087135	Mus musculus cto
gb_htg:AC027236	170186	5.9e+04	75.96	39.00	170186	AC027236	Homo sapiens chromd	gb_htg:AL590614	39.00	74.06	7.5e+04	211033	AL590614	Mus musculus chr
gb_in:AC009250	170362	5.9e+04	75.95	39.00	170362	AC009250	Drosophila melanoga	gb_htg:AL353668	39.00	74.02	7.5e+04	211967	AL353668	Homo sapiens cto
gb_htg:AC092045	171321	5.9e+04	75.90	39.00	171321	AC092045	Homo sapiens chromd	gb_htg:AC073883	39.00	73.97	7.5e+04	212987	AC073883	Mus musculus chr
em_htg:AL442004	172033	5.9e+04	75.86	39.00	172033	AC024148	Homo sapiens chromd	gb_htg:AC093339	39.00	73.88	7.6e+04	213552	AC093339	Mus musculus cto
gb_htg:AL442004	172219	5.9e+04	75.85	39.00	172219	AL442004	Homo sapiens chromd	gb_htg:AC090293	39.00	73.85	7.6e+04	216016	AC090293	Mus musculus chr
gb_htg:AL591127	172385	5.9e+04	75.85	39.00	172385	AL591127	Mus musculus chromd	gb_htg:AL592547	39.00	73.76	7.7e+04	218205	AL592547	Mus musculus chr
gb_ro:CNS07EGP	172751	6.0e+04	75.83	39.00	172751	AL591826	BAC 13C18 of libra	gb_htg:AC034265	39.00	73.76	7.7e+04	218270	AC034265	Mus musculus cto
gb_htg:AC013729	173327	6.0e+04	75.80	39.00	173327	AC013729	Homo sapiens chromd	gb_htg:AC083890	39.00	73.73	7.8e+04	219038	AC083890	Mus musculus chr
gb_pr:AP002360	173379	6.0e+04	75.79	39.00	173379	AP002360	Homo sapiens genom	gb_htg:AC087556	39.00	73.67	7.8e+04	220460	AC087556	Mus musculus chr
gb_pr:CNS05OTC	173389	6.0e+04	75.79	39.00	173389	AL355103	Homo sapiens chromd	gb_ba:AP002555	39.00	73.58	7.9e+04	222605	AP002555	Escherichia coli
gb_htg:AC068960	173975	6.0e+04	75.76	39.00	173975	AC068960	Homo sapiens chromd	gb_htg:AL358779	39.00	73.49	8.0e+04	225022	AL358779	Homo sapiens chr
gb_pr:AC011841	174588	6.0e+04	75.73	39.00	174588	AC011841	Homo sapiens chromd	gb_htg:AC027281	39.00	73.48	8.0e+04	225157	AC027281	Homo sapiens chr
gb_htg:AC023889	174612	6.0e+04	75.73	39.00	174612	AC023889	Homo sapiens chromd	gb_htg:AC073680	39.00	73.45	8.0e+04	226013	AC073680	Mus musculus cto
gb_htg:AC073748	175371	6.1e+04	75.69	39.00	175371	AC073748	Mus musculus clone	gb_htg:AC091287	39.00	73.39	8.1e+04	227518	AC091287	Mus musculus cto
gb_htg:AC027619	175399	6.1e+04	75.69	39.00	175399	AC027619	Homo sapiens chromd	gb_htg:AC090652	39.00	73.37	8.1e+04	228159	AC090652	Mus musculus cto
em_htg:hum:AC012036	177341	6.1e+04	75.60	39.00	177341	AC012036	Homo sapiens chromd	gb_htg:AC090122	39.00	73.30	8.2e+04	229774	AC090122	Mus musculus cto
gb_htg:AC032013	177421	6.1e+04	75.59	39.00	177421	AC032013	Mus musculus chromd	gb_vl:AF232689	39.00	73.30	8.2e+04	229996	AF232689	Rat cytomegalovi
gb_htg:AC079443	177809	6.1e+04	75.57	39.00	177809	AC079443	Mus musculus chromd	gb_htg:AC073691	39.00	73.24	8.3e+04	231443	AC073691	Mus musculus cto
gb_htg:AC025163	177876	6.2e+04	75.57	39.00	177876	AC025163	Homo sapiens chromd	gb_htg:AC079638	39.00	73.00	8.5e+04	237861	AC079638	Mus musculus cto
gb_htg:AC080155	177924	6.2e+04	75.57	39.00	177924	AC080155	Rattus norvegicus c	gb_htg:AC092254	39.00	72.97	8.6e+04	238637	AC092254	Mus musculus cto
gb_htg:AC087869	178032	6.2e+04	75.56	39.00	178032	AC087869	Mus musculus clone	gb_htg:AC044892	39.00	72.91	8.6e+04	240294	AC044892	Homo sapiens chr
gb_htg:AC026626	178953	6.2e+04	75.51	39.00	178953	AC026626	Homo sapiens chromd	gb_htg:AL360154	39.00	72.90	8.6e+04	240579	AL360154	Homo sapiens chr
gb_htg:AC010178	179070	6.2e+04	75.51	39.00	179070	AC010178	Homo sapiens chromd	gb_htg:AL589661	39.00	72.87	8.7e+04	241226	AL589661	Mus musculus chr
gb_pr:AC016643	179969	6.2e+04	75.46	39.00	179969	AC016643	Homo sapiens chromd	gb_htg:AC025909	39.00	72.63	8.9e+04	247899	AC025909	Mus musculus chr
gb_htg:AP003766	180865	6.3e+04	75.42	39.00	180865	AP003766	Oryza sativa chromd	gb_in:AE003653	39.00	72.58	9.0e+04	249469	AE003653	Drosophila melan
gb_htg:AC062006	181062	6.3e+04	75.41	39.00	181062	AC062006	Homo sapiens chromd	gb_htg:AC073677	39.00	72.50	9.1e+04	251579	AC073677	Mus musculus cto
gb_htg:AC072057	181452	6.3e+04	75.39	39.00	181452	AC072057	Homo sapiens chromd	gb_ba:AP000007	39.00	72.50	9.1e+04	251579	AP000007	Pyrococcus horik
gb_htg:AC025878	182204	6.3e+04	75.36	39.00	182204	AC025878	Homo sapiens clone	gb_htg:AL513124	39.00	72.43	9.1e+04	253505	AL513124	Homo sapiens chr
gb_htg:AL592464	182766	6.3e+04	75.33	39.00	182766	AL592464	Homo sapiens chromd	gb_htg:AC087228	39.00	72.17	9.5e+04	261141	AC087228	Homo sapiens chr
gb_htg:AC040980	183150	6.4e+04	75.31	39.00	183150	AC040980	Homo sapiens chromd	gb_htg:AC087129	39.00	72.02	9.6e+04	265537	AC087129	Mus musculus chr
gb_htg:AP002000	183197	6.4e+04	75.31	39.00	183197	AP002000	Homo sapiens chromd	gb_htg:AC068493	39.00	71.99	9.7e+04	266574	AC068493	Mus musculus cto
gb_htg:AC068828	183487	6.4e+04	75.29	39.00	183487	AC068828	Homo sapiens clone	gb_htg:AC092180	39.00	71.91	9.8e+04	269081	AC092180	Homo sapiens chr
gb_pr:AC020709	184104	6.4e+04	75.26	39.00	184104	AC020709	Homo sapiens clone	gb_htg:AC092098	39.00	71.81	9.9e+04	272120	AC092098	Mus musculus chr
gb_htg:AC025861	185143	6.4e+04	75.21	39.00	185143	AC025861	Homo sapiens clone	gb_htg						

gb_hgt:AL356291	39.00	70.44	1.2e+05	317436	! AL356291 Homo sapiens chromid	gb_ba:AE004227	-	38.00	97.17	3.9e+03	10177	! AE004227 Vibrio cholerae
gb_hgt:AL356291	39.00	70.26	1.2e+05	324165	! AL356291 Homo sapiens chromid	gb_ba:AE001922	+	38.00	96.97	4.0e+03	10410	! AE001922 Deinococcus radi
gb_ba:AP003005	39.00	70.03	1.2e+05	323635	! AP003005 Mesorhizobium loti	gb_ba:AE0004595	+	38.00	96.94	4.0e+03	10448	! AE0004595 Pseudomonas aerug
gb_ba:AP003010	39.00	69.81	1.3e+05	340857	! AP003010 Mesorhizobium loti	gb_ba:AE0004618	+	38.00	96.43	4.3e+03	11060	! AE0004618 Pseudomonas aerug
gb_ba:AP002095	39.00	69.66	1.3e+05	348897	! AP002095 Mesorhizobium loti	gb_pl:SPAC32A11	-	38.00	96.17	4.4e+03	11401	! Z69796 S.pombe chromosome
gb_ba:AP003007	39.00	69.62	1.3e+05	348411	! AP003007 Mesorhizobium loti	gb_ba:AE0005669	-	38.00	96.01	4.5e+03	11599	! AE0005669 Escherichia coli
gb_ba:AE0005019	38.50	99.04	3.1e+03	196646	! AE0005019 Halobacterium sp. NH	gb_ba:AE000148	+	38.00	95.94	4.6e+03	11692	! AE000148 Escherichia coli
gb_ba:AE0073146	38.50	72.83	8.7e+04	196646	! AE0073146 Homo sapiens chromo	gb_ba:AE0005221	+	38.00	95.83	4.6e+03	11845	! AE005221 Escherichia coli
gb_sts:G28085	38.00	125.10	108.24	433	! G28085 human STS SHGC-33555, se	gb_hgt:AC020143	+	38.00	95.82	4.6e+03	12054	! AC020143 Streptomyces gris
gb_ba:BU01793	38.00	122.94	142.87	553	! BU01793 Borrelia burgdorferi str	gb_ba:AB011413	+	38.00	95.66	4.7e+03	12070	! AB011413 Drosophila melano
gb_ba:BOR220SPC	38.00	122.72	146.98	567	! L42889 Borrelia burgdorferi (st	gb_ba:AE000769	+	38.00	95.31	4.9e+03	12563	! AE000769 Aquifex aeolicus
gb_ba:BG00SPC1	38.00	121.87	163.86	624	! X81525 B.gariniia (strain Tis1)	gb_ba:AE0005970	+	38.00	95.23	5.0e+03	12680	! AE0005970 Caulobacter cresc
gb_ba:AE0051422	38.00	121.42	173.73	657	! AE0051422 Giardia intestinalis c	gb_ba:AE0004759	+	38.00	94.15	5.7e+03	14309	! AE0004759 Pseudomonas aerug
gb_sts:G73117	38.00	121.28	176.73	667	! G73117 MARC 7899-7900:996688143	gb_ba:AE0006848	+	38.00	93.88	5.9e+03	14759	! AE006848 Sulfolobus solfat
gb_ro:MUSTGG4014	38.00	120.83	187.29	702	! M27660 Mus musculus rearranged	gb_pat:AR093181	+	38.00	93.68	6.1e+03	15101	! AR093181 Sequence 14 from
gb_pat:105395	38.00	119.97	209.24	774	! I05395 Sequence 2 from Patent B	gb_pat:A36843	+	38.00	93.68	6.1e+03	15108	! A26843 Lelystad Agent-spec
gb_pr:HSY13782	38.00	119.97	209.24	774	! I13782 Homo sapiens mRNA for te	gb_pat:ARL37155	+	38.00	93.68	6.1e+03	15108	! AR137155 Sequence 1 from p
gb_pat:AX056449	38.00	119.43	224.02	822	! AX056449 Sequence 93 from Patent	gb_pat:1A40237	+	38.00	93.68	6.1e+03	15108	! I40237 Sequence 1 from pat
gb_pat:AX1390530	38.00	119.42	224.33	823	! AX1390530 Sequence 1 from Patent	gb_vi:LEXPOLYENV	+	38.00	93.67	6.1e+03	15111	! M96262 Lelystad virus, com
gb_hgt:AC037791	38.00	118.83	242.04	880	! AC037791 Giardia intestinalis c	gb_ba:AE0000934	+	38.00	93.67	6.1e+03	15124	! AE000934 Methanobacterium
gb_hgt:AC038802	38.00	118.44	254.57	920	! AC038802 Giardia intestinalis c	gb_hgt:AC090678	+	38.00	93.60	6.2e+03	15239	! AC090678 Homo sapiens chro
gb_hgt:AC076604	38.00	118.42	255.19	922	! AC076604 Giardia intestinalis c	gb_ba:AE005026	-	38.00	93.53	6.2e+03	15355	! AE005026 Halobacterium sp.
gb_hgt:AC080544	38.00	118.31	258.65	933	! AC080544 Giardia intestinalis c	gb_ba:AE0000508	-	38.00	93.44	6.3e+03	15518	! AE000508 Escherichia coli
gb_ba:BORWNP3A	38.00	118.06	267.16	960	! L24911 Borrelia hermsii outer m	gb_ba:AE0004094	-	38.00	93.08	6.6e+03	16164	! AE004094 Vibrio cholerae c
gb_hgt:AC062559	38.00	118.05	267.48	961	! AC062559 Giardia intestinalis c	gb_hgt:AC0017330	+	38.00	92.80	6.8e+03	16673	! AC0017330 Mycobacterium tub
gb_hgt:AC051183	38.00	117.91	272.22	981	! AC051183 Giardia intestinalis c	gb_ba:AE0007029	+	38.00	92.16	7.4e+03	17286	! AE007029 Mycobacterium tub
gb_in:AF217397	38.00	117.87	273.80	986	! AF217397 Drosophila melanogaster	gb_ro:MMTBL1	+	38.00	92.48	7.7e+03	17934	! F38006 Mus musculus partia
gb_hgt:AC051094	38.00	116.82	313.08	1104	! AC051094 Giardia intestinalis	gb_ba:BSPHRZ	-	38.00	91.25	8.3e+03	19861	! Z80360 B.subtilis thrz dow
gb_pat:AX133781	38.00	115.65	363.74	1260	! AX133781 Sequence 1 from Patent	gb_pl:SPAC750	+	38.00	89.47	1.0e+04	24290	! AL390095 S.pombe chromosom
gb_ro:AF329484	38.00	115.20	385.76	1327	! AF329484 Mus musculus serine/Ar	gb_ba:SCDL2A	-	38.00	87.41	1.4e+04	30657	! AL357524 Streptomyces coel
gb_in:DROGABARBS	38.00	114.16	440.30	1491	! L17436 Drosophila melanogaster	gb_ba:SCBE4A	-	38.00	86.74	1.5e+04	33084	! AL138662 Streptomyces coel
gb_in:SE62717	38.00	114.16	440.30	1491	! SE62717 LCCH3-putative gamma-am	gb_ba:MTGCV01B2	-	38.00	86.01	1.6e+04	35938	! Z95554 Mycobacterium tuber
gb_pat:AR068511	38.00	114.16	440.30	1491	! AR068511 Sequence 1 from patent	gb_hgt:AC014499	+	38.00	85.76	1.7e+04	36948	! AC014499 Drosophila melano
gb_pat:117317	38.00	114.16	440.30	1491	! I17317 Sequence 3 from patent	gb_pl:SPAC56P8	-	38.00	85.73	1.7e+04	37093	! Z69728 S.pombe chromosome
gb_pat:AX141289	38.00	113.85	458.44	1545	! AX141289 Sequence 11 from Patent	gb_hgt:AC080114	-	38.00	85.66	1.7e+04	37367	! AC080114 Homo sapiens chro
gb_ro:AF329483	38.00	113.80	461.47	1554	! AF329483 Homo sapiens serine/Ar	gb_pl:SPAC1348	-	38.00	85.24	1.8e+04	39187	! AL358912 S.pombe chromosom
gb_ro:AF329483	38.00	113.78	462.82	1558	! AF329483 Rattus norvegicus pro	gb_pl:SPAC977	+	38.00	84.88	1.9e+04	40815	! AL371330 S.pombe chromosom
gb_in:SE62717	38.00	113.43	483.75	1620	! D17333 E. Coli thil gene, comp	gb_in:CBRG45N02	-	38.00	84.48	2.0e+04	42726	! AC084649 Caenorhabditis br
gb_pr:HSZA293345	38.00	113.27	493.95	1650	! ZJ293345 Homo sapiens partia	gb_pr:AC007206	+	38.00	84.48	2.0e+04	42732	! AC007206 Homo sapiens chro
gb_pl:AF367339	38.00	112.01	580.40	1902	! AF367339 Arabidopsis thaliana	gb_hgt:AC019872	-	38.00	84.08	2.1e+04	44671	! AC019872 Drosophila melano
gb_pl:AF105141	38.00	111.71	603.31	1968	! AF105141 Brassica napus chromo	gb_hgt:AC019806	-	38.00	84.02	2.1e+04	44986	! AC019806 Drosophila melano
gb_in:LMSAPIGN	38.00	111.68	605.40	1974	! Z46969 L.mexicana lmsap1 gene	gb_hgt:AC091265	-	38.00	83.84	2.2e+04	45912	! AC091265 Mus musculus clon
gb_pl:CPGCRASA	38.00	111.44	624.57	2029	! M34606 C.neoformans orotidine	gb_ph:AL136168	-	38.00	83.66	2.2e+04	46851	! AL136168 Human DNA sequen
gb_pr:AC056371	38.00	110.73	684.28	2199	! AR056371 Macaca fascicularis b	gb_pr:AF022214	+	38.00	83.24	2.3e+04	49136	! AF022214 Mycobacteriophage
gb_ba:BACR7NAXSA	38.00	110.23	729.94	2325	! M36593 B.subtilis threonyl-TRN	gb_pat:AR014185	+	38.00	83.22	2.3e+04	49272	! AR014185 Sequence 1 from p
gb_pr:HDMP2C9X01	38.00	109.81	768.28	2438	! L16877 Homo sapiens cytochrome	gb_hgt:AC017572	-	38.00	83.03	2.4e+04	50319	! AC017572 Drosophila melano
gb_pr:AB055297	38.00	109.39	812.02	2557	! AB055297 Macaca fascicularis b	gb_pl:SPBPB2B2	-	38.00	82.79	2.5e+04	53715	! AL512522 S.pombe chromosom
gb_ro:HSAL17866	38.00	109.25	826.45	2597	! Y17866 Homo sapiens mRNA for t	gb_hgt:AC007800	-	38.00	82.61	2.5e+04	52737	! AC007800 Homo sapiens chro
gb_ro:MUSPTPE	38.00	108.50	906.99	2827	! D83844 Mouse mRNA for protein	gb_hgt:AC0079961	-	38.00	82.62	2.7e+04	56154	! AC007961 Homo sapiens chro
gb_in:AF148553	38.00	108.05	964.98	2977	! AF148553 Thioabacillus denitrif	gb_ro:D78344	-	38.00	81.53	2.9e+04	59641	! D78344 Mouse DNA for Ig g
gb_in:LMSAP2GN	38.00	107.93	978.97	3015	! Z46970 L.mexicana lmsap2 gene	gb_hgt:AC022549	-	38.00	81.46	2.9e+04	60058	! AC022549 Mus musculus clon
gb_ba:AF1155058	38.00	107.90	983.76	3028	! AF155058 Israeli tick typhus r	gb_hgt:AC014595	-	38.00	80.64	3.2e+04	65957	! AC014595 Drosophila melano
gb_ba:AF269408	38.00	107.71	1.0e+03	3093	! AF269408 Staphylococcus epide	gb_hgt:AC083985	-	38.00	80.21	3.4e+04	69227	! AC083985 Homo sapiens chro
gb_pat:AX1434728	38.00	107.71	1.0e+03	3093	! AX1434728 Sequence 3450 from Pat	gb_pl:NCBFF18	-	38.00	79.97	3.5e+04	71088	! AL389891 Neurospora crassa
gb_pr:ASZA922757	38.00	106.63	1.2e+03	3495	! AF292757 Homo sapiens TUBB1 ge	gb_hgt:AC069119	-	38.00	79.71	3.6e+04	73256	! AC069119 Homo sapiens chro
gb_ba:RCU01133	38.00	106.61	1.2e+03	3500	! U01133 Rickettsia conorii Mal	gb_hgt:AC0614256	-	38.00	79.29	3.8e+04	76748	! AC0614256 Drosophila melano
gb_ba:AF270407	38.00	104.90	1.4e+03	4250	! AF270407 Staphylococcus epide	gb_pl:NCB24P7	-	38.00	79.18	3.9e+04	77715	! AL389890 Neurospora crassa
gb_pat:AX145725	38.00	104.90	1.4e+03	4250	! AX145725 Sequence 4447 from Pat	gb_pl:NC03G11	+	38.00	79.28	4.4e+04	86075	! AL513443 Neurospora crassa
gb_ro:MMU35368	38.00	102.74	1.9e+03	5425	! U35368 Mus musculus epsilon ty	gb_pr:AC0008585	-	38.00	78.27	4.4e+04	86152	! AC008585 Homo sapiens chro
gb_pl:BEUPPOA	38.00	102.28	2.0e+03	5711	! L27746 Neurospora crassa N-ace	gb_in:CEX102A5C	-	38.00	77.76	4.7e+04	91286	! AL031627 Arabidopsis el
gb_ba:BMU17226	38.00	102.02	2.1e+03	5895	! U17226 Rhizobium meliloti beta	gb_hgt:AC007166	+	38.00	77.75	4.7e+04	91382	! AC007166 Arabidopsis thali
gb_in:LMLPKORF	38.00	101.27	2.3e+03	6404	! Z95867 Leishmania mexicana imp	gb_pr:AC0004254	-	38.00	77.52	4.8e+04	93821	! AC004254 Homo sapiens BAC
gb_in:AX028181	38.00	101.05	2.4e+03	6568	! AX028181 Sequence 1 from Patent	gb_pr:HSBC17A96	-	38.00	77.49	4.8e+04	94074	! AL450224 Homo sapiens chro
gb_pr:AB0202311	38.00	101.05	2.4e+03	6568	! AB0202311 Human mRNA for KIA03	gb_pr:AC073133	-	38.00	77.38	4.9e+04	95280	! AC073133 Homo sapiens chro
gb_ba:AF012886	38.00	100.79	2.4e+03	6757	! AF012886 Buchnera aphidicola U	gb_hgt:HSU600024	-	38.00	77.28	4.9e+04	95387	! AL080247 Homo sapiens chro
gb_pr:AB033019	38.00	100.33	2.6e+03	7118	! AB033019 Homo sapiens mRNA for	gb_hgt:AC017943	+	38.00	77.12	5.0e+04	97538	! AC017943 Drosophila melano
gb_vi:AF145896	38.00	99.81	2.8e+03	7555	! AF145896 Camberwell virus poly	gb_pr:AP000072	-	38.00	76.95	5.2e+04	100000	! AF000072 Homo sapiens gen
gb_in:AF254867	38.00	99.26	3.0e+03	8035	! AF254867 Drosophila melanogast	gb_hgt:AC020372	-	38.00	76.95	5.2e+04	100002	! AC020372 Drosophila melan
gb_ba:U39680	38.00	98.84	3.1e+03	8430	! U39680 Mycoplasma genitalium s	gb_pr:HSWMC3W36A	-	38.00	76.93	5.2e+04	100267	! U89337 Homo sapiens HLA c
gb_ba:AE006023	38.00	98.64	3.2e+03	8617	! AE006023 Caulobacter crescentu	gb_pr:AC003962	-	38.00	76.83	5.2e+04	101370	! AC003962 Homo sapiens chr
gb_ro:AE006115	38.00	98.37	3.3e+03	8891	! AE006115 Pasteurella multocida	gb_hgt:AC025862	-	38.00	76.31	5.6e+04	107558	! AC025862 Homo sapiens clo
gb_ov:XLRLP14	38.00	98.14	3.4e+03	9124	! AF05025 Xenopus laevis gene for	gb_hgt:AC009839	-	38.00	76.13	5.7e+04	109750	! AC009839 Drosophila melan
gb_ro:AF083064	38.00	97.18	3.9e+03	10172	! AF083064 Mus musculus Paxillin	gb_hgt:AC068947	0	38.00	76.11	5.7e+04	110000	! AC068947 Mus musculus clo

gb_htg:AC091288_1	-	38.00	76.11	5.7e+04	110000	Continuation (2 of 4) of AC091288_1	gb_htg:AC008336	+	38.00	72.61	8.9e+04	163378	AC008336	Drosophila melan
gb_htg:AC091288_2	-	38.00	76.11	5.7e+04	110000	Continuation (3 of 4) of AC091288_1	gb_htg:AL357122	-	38.00	72.59	9.0e+04	163338	AL357122	Homo sapiens chr
gb_htg:LMFCHR15_5	+	38.00	76.11	5.7e+04	110000	Continuation (6 of 8) of LMFCHR15_5	gb_htg:AC025385	-	38.00	72.58	9.0e+04	163337	AC025385	Homo sapiens chr
gb_htg:LMFCHR15_6	+	38.00	76.11	5.7e+04	110000	Continuation (7 of 8) of LMFCHR15_6	gb_htg:AC092531	-	38.00	72.57	9.0e+04	164032	AC092531	Rattus norvegicus
gb_htg:LMFCHR16_08	+	38.00	76.11	5.7e+04	110000	Continuation (9 of 11) of LMFCHR16_08	gb_htg:AL390855	-	38.00	72.54	9.0e+04	164736	AL390855	Homo sapiens chr
gb_htg:LMFCHR36_05	-	38.00	76.11	5.7e+04	110000	Continuation (6 of 36) of LMFCHR36_05	gb_pr:AC011448	+	38.00	72.52	9.1e+04	165122	AC011448	Homo sapiens chr
gb_htg:LMFCHR36_06	-	38.00	76.11	5.7e+04	110000	Continuation (7 of 36) of LMFCHR36_06	gb_htg:AC068858	+	38.00	72.51	9.1e+04	165274	AC068858	Homo sapiens chr
gb_htg:LMFCHR36_07	-	38.00	76.11	5.7e+04	110000	Continuation (8 of 36) of LMFCHR36_07	gb_htg:AC016663	+	38.00	72.50	9.1e+04	165415	AC016663	Homo sapiens chr
gb_htg:LMFCHR36_24	-	38.00	76.11	5.7e+04	110000	Continuation (25 of 36) of LMFCHR36_24	gb_htg:AC013552	+	38.00	72.47	9.1e+04	166065	AC013552	Homo sapiens chr
gb_htg:LMFCHR36_33	+	38.00	76.11	5.7e+04	110000	Continuation (34 of 36) of LMFCHR36_33	gb_htg:AC083894	+	38.00	72.45	9.1e+04	166407	AC083894	Homo sapiens chr
gb_pr:HS175E3	+	38.00	76.11	5.7e+04	110042	295113 Human DNA sequence fr	gb_ro:AF332859S2	+	38.00	72.44	9.1e+04	166617	AF332862	Mus musculus dia
gb_htg:AL139143	+	38.00	75.96	5.9e+04	111862	AL139143 Homo sapiens chromo	gb_htg:AC025505	+	38.00	72.43	9.1e+04	166695	AC025505	Homo sapiens chr
gb_pr:AP003373	+	38.00	75.91	5.9e+04	112509	AP003373 Oryza sativa chromo	gb_htg:AC021103	+	38.00	72.43	9.1e+04	166695	AC021103	Homo sapiens chr
gb_pr:AL512442	-	38.00	75.80	6.0e+04	113884	AL512442 Human DNA sequence	gb_htg:AC023340	+	38.00	72.40	9.2e+04	167339	AC023340	Homo sapiens chr
gb_pr:AL591025	-	38.00	75.76	6.0e+04	114477	AL591025 Human DNA sequence	gb_htg:AC022785	+	38.00	72.36	9.2e+04	168032	AC022785	Homo sapiens chr
gb_pr:HS173D1	+	38.00	75.54	6.2e+04	117338	AL031984 Human DNA sequence	gb_htg:AC021971	-	38.00	72.36	9.2e+04	168124	AC021971	Homo sapiens chr
gb_htg:AC023193	+	38.00	75.52	6.2e+04	117542	AC023193 Mus musculus chromo	gb_pr:AC018719	-	38.00	72.35	9.2e+04	168210	AC018719	Homo sapiens chr
gb_htg:AC020548	+	38.00	75.45	6.2e+04	118492	AC020548 Homo sapiens chromo	gb_htg:AC046148	-	38.00	72.34	9.3e+04	168524	AC046148	Mus musculus chr
gb_pr:AF275345	-	38.00	75.43	6.3e+04	118813	AF275345 Lycopersicon escul	gb_in:AC007471	-	38.00	72.33	9.3e+04	168601	AC007471	Drosophila melan
gb_htg:AC010112	-	38.00	75.19	6.5e+04	122073	AC010112 Drosophila melanog	gb_ba:AP002569	-	38.00	72.32	9.3e+04	168764	AP002569	Escherichia coli
gb_htg:AL161737	+	38.00	75.16	6.5e+04	122493	AL161737 Homo sapiens chromo	gb_pr:HS490024	+	38.00	72.32	9.3e+04	168778	AL117338	Human DNA sequen
gb_pr:HSBCL17A99	-	38.00	75.09	6.5e+04	123454	AL450226 Homo sapiens chromo	gb_htg:AC024288	+	38.00	72.28	9.3e+04	169540	AC024288	Homo sapiens chr
gb_pr:HSJ958M22	-	38.00	75.04	6.6e+04	124114	AL118510 Human DNA sequence	gb_htg:AC087257	+	38.00	72.25	9.4e+04	170116	AC087257	Homo sapiens chr
gb_pr:AL450267	+	38.00	74.97	6.6e+04	125110	AL450267 Human DNA sequence	gb_in:AC008094	+	38.00	72.23	9.4e+04	170498	AC008094	Drosophila melan
gb_htg:AL139255	+	38.00	74.97	6.6e+04	125114	AL139255 Homo sapiens chromo	gb_htg:AC090965	+	38.00	72.22	9.4e+04	170745	AC090965	Papio cynocephal
gb_pr:HS1018K9	-	38.00	74.94	6.7e+04	125500	AL031726 Human DNA sequence	gb_in:AC007757	+	38.00	72.21	9.4e+04	170939	AC007757	Drosophila melan
gb_pr:AL161424	-	38.00	74.88	6.7e+04	126427	AL161424 Human DNA sequence	gb_htg:AC007357	+	38.00	72.18	9.4e+04	171589	AC007357	Homo sapiens chr
gb_htg:H0201G08	+	38.00	74.79	6.8e+04	127673	AL442118 Oryza sativa chromo	gb_htg:AL359175	-	38.00	72.17	9.5e+04	171598	AL359175	Homo sapiens chr
gb_pr:HSBQ34F7	-	38.00	74.65	6.9e+04	129811	AL049547 Human DNA sequence	gb_pr:AC006337	+	38.00	72.14	9.5e+04	172291	AC006337	Homo sapiens BAC
gb_htg:AC009383	-	38.00	74.35	7.2e+04	13261	AC009383 Drosophila melanog	gb_htg:AL450998	-	38.00	72.13	9.5e+04	172431	AL450998	Homo sapiens chr
gb_pr:AL512662	+	38.00	74.34	7.2e+04	134411	AL512662 Homo sapiens chromo	gb_htg:AL133513	-	38.00	72.10	9.5e+04	173154	AL133513	Homo sapiens chr
gb_pr:AC008496	-	38.00	74.27	7.3e+04	135429	AC008496 Homo sapiens chromo	gb_htg:AC092818	-	38.00	72.06	9.6e+04	173762	AC092818	Homo sapiens chr
gb_pr:AC004843	-	38.00	74.19	7.3e+04	136655	AC004843 Homo sapiens PAC cl	gb_in:AC006562	-	38.00	72.06	9.6e+04	173781	AC006562	Drosophila melan
gb_ba:SC082664	+	38.00	73.99	7.5e+04	139818	U82664 Escherichia coli mini	gb_htg:AL358160	-	38.00	72.06	9.6e+04	173852	AL358160	Homo sapiens chr
gb_ro:AC006447	+	38.00	73.95	7.6e+04	140452	AC006447 Mus musculus chromo	gb_htg:AC091671	-	38.00	72.06	9.6e+04	173957	AC091671	Papio cynocephal
gb_htg:AC010919	+	38.00	73.93	7.6e+04	140685	AC010919 Drosophila melanog	gb_pr:AC032263	+	38.00	72.04	9.6e+04	174176	AC032263	Homo sapiens, cl
gb_htg:AL591932	-	38.00	73.90	7.6e+04	141199	AL591932 Homo sapiens chromo	gb_htg:AL596264	-	38.00	72.02	9.6e+04	174641	AL596264	Mus musculus chr
gb_pr:HSJ543J19	+	38.00	73.85	7.7e+04	142094	AL109840 Human DNA sequence	gb_in:AC018479	-	38.00	72.02	9.6e+04	174673	AC018479	Drosophila melan
gb_ro:AC006404	+	38.00	73.80	7.7e+04	142776	AC006404 Mus musculus chromo	gb_htg:AC016106	-	38.00	72.01	9.6e+04	174823	AC016106	Homo sapiens chr
gb_htg:AL359672	-	38.00	73.78	7.7e+04	143087	AL359672 Homo sapiens chromo	gb_pr:AC009069	-	38.00	72.01	9.7e+04	174873	AC009069	Homo sapiens chr
gb_htg:AC022662	+	38.00	73.66	7.8e+04	143424	AC022662 Homo sapiens clone	gb_htg:AC026282	+	38.00	72.00	9.7e+04	174992	AC026282	Homo sapiens chr
gb_pr:AC011712	+	38.00	73.62	7.8e+04	145165	AC011712 Homo sapiens chromo	gb_htg:AC058783	-	38.00	72.00	9.7e+04	175066	AC058783	Homo sapiens chr
gb_htg:AC016588	-	38.00	73.62	7.9e+04	145684	AC016588 Homo sapiens chromo	gb_ro:AC091536	-	38.00	71.98	9.7e+04	175512	AC091536	Rattus norvegicus
gb_pr:AC022234	-	38.00	73.53	8.0e+04	147191	AC022234 Homo sapiens chromo	gb_in:AC007186	-	38.00	71.94	9.7e+04	176306	AC007186	Drosophila melan
gb_in:AC008234	-	38.00	73.43	8.1e+04	148847	AC008234 Drosophila melanog	gb_pr:HS402611	-	38.00	71.89	9.8e+04	177241	AL022328	Human DNA sequen
gb_pr:AC008033	-	38.00	73.41	8.1e+04	149180	AC008033 Homo sapiens 12q15-	gb_htg:AC008494	-	38.00	71.89	9.8e+04	177281	AC008494	Homo sapiens chr
gb_htg:AC016286	+	38.00	73.40	8.1e+04	149415	AC016286 Homo sapiens clone	gb_htg:AC097775	-	38.00	71.88	9.8e+04	177363	AC097775	Homo sapiens chr
gb_ro:AL36084	+	38.00	73.37	8.1e+04	149930	AL136084 Human DNA sequence	gb_htg:AL591862	+	38.00	71.86	9.8e+04	177477	AL591862	Homo sapiens chr
gb_htg:AC074047	+	38.00	73.35	8.2e+04	150351	AC074047 Mus musculus 7 BAC	gb_pr:AP000925	+	38.00	71.86	9.8e+04	177830	AP000925	Homo sapiens gen
gb_htg:AC021762	+	38.00	73.32	8.2e+04	150817	AC021762 Homo sapiens chromo	gb_htg:AL390776	-	38.00	71.83	9.9e+04	178431	AL390776	Homo sapiens chr
gb_htg:AC092502	+	38.00	73.30	8.2e+04	151198	AC092502 Homo sapiens chromo	gb_pr:AC069020	-	38.00	71.83	9.9e+04	178424	AC069020	Homo sapiens chr
gb_htg:AC069315	-	38.00	73.23	8.3e+04	152394	AC069315 Homo sapiens chromo	gb_htg:AC091782	-	38.00	71.83	9.9e+04	178436	AC091782	Mus musculus chr
gb_htg:AL18943	-	38.00	73.20	8.3e+04	152779	AL18943 Homo sapiens clone	gb_htg:AC032028	-	38.00	71.81	9.9e+04	178803	AC032028	Homo sapiens chr
gb_htg:AC010344	+	38.00	73.20	8.3e+04	152829	AC010344 Homo sapiens chromo	gb_htg:AL390786	-	38.00	71.76	1.0e+05	179927	AL390786	Homo sapiens chr
gb_htg:AC024031	-	38.00	73.16	8.4e+04	153528	AC024031 Homo sapiens chromo	gb_ba:AF232012S1	-	38.00	71.75	1.0e+05	180000	AF232012	Bradyrhizobium j
gb_htg:AC023953	+	38.00	73.14	8.4e+04	153937	AC023953 Homo sapiens chromo	gb_htg:AC009831	-	38.00	71.75	1.0e+05	180041	AC009831	Homo sapiens chr
gb_htg:AC019027	+	38.00	73.14	8.4e+04	153951	AC019027 Homo sapiens clone	gb_htg:AC024276	+	38.00	71.73	1.0e+05	180368	AC024276	Homo sapiens chr
gb_htg:AC024885	+	38.00	73.02	8.5e+04	155718	AC024885 Homo sapiens chromo	gb_htg:AC010776	+	38.00	71.68	1.0e+05	181431	AC010776	Homo sapiens chr
gb_htg:AC015539	+	38.00	73.02	8.5e+04	155979	AC015539 Homo sapiens chromo	gb_ro:AC069014	+	38.00	71.68	1.0e+05	181557	AC069014	Mus musculus il
gb_htg:AC074359	+	38.00	72.95	8.6e+04	157225	AC074359 Mus musculus chromo	gb_htg:AL591425	-	38.00	71.66	1.0e+05	181810	AL591425	Mus musculus chr
gb_htg:AC007606	+	38.00	72.94	8.6e+04	157375	AC007606 Homo sapiens chromo	gb_in:AC008093	+	38.00	71.66	1.0e+05	181950	AC008093	Drosophila melan
gb_htg:AC020598	-	38.00	72.85	8.7e+04	158983	AC020598 Homo sapiens chromo	gb_htg:AC015932	-	38.00	71.63	1.0e+05	182074	AC015932	Mus musculus chr
gb_htg:AC013649	+	38.00	72.85	8.7e+04	159082	AC013649 Homo sapiens clone	gb_htg:AC019180	+	38.00	71.63	1.0e+05	182504	AC019180	Homo sapiens chr
gb_pr:AL161781	+	38.00	72.85	8.7e+04	159096	AL161781 Human DNA sequence	gb_htg:AC022374	-	38.00	71.62	1.0e+05	182706	AC022374	Homo sapiens chr
gb_htg:AC012585	+	38.00	72.81	8.7e+04	159668	AC012585 Homo sapiens clone	gb_htg:AC090964	-	38.00	71.62	1.0e+05	182766	AC090964	Papio cynocephal
gb_pr:AC007179	+	38.00	72.81	8.7e+04	159791	AC007179 Homo sapiens BAC cl	gb_htg:AC037435	+	38.00	71.56	1.0e+05	183937	AC037435	Mus musculus chr
gb_pr:AC074344	-	38.00	72.78	8.8e+04	160185	AC074344 Homo sapiens clone	gb_htg:AC024967	+	38.00	71.56	1.0e+05	183978	AC024967	Homo sapiens chr
gb_htg:AC067928	-	38.00	72.75	8.8e+04	160750	AC067928 Homo sapiens clone	gb_htg:AC026277	+	38.00	71.54	1.0e+05	184477	AC026277	Homo sapiens chr
gb_pr:CN057ED0	+	38.00	72.75	8.8e+04	160750	AL445384 Human chromosome 14	gb_htg:AP002388	+	38.00	71.50	1.0e+05	185147	AP002388	Homo sapiens chr
gb_htg:AL591936	+	38.00	72.71	8.8e+04	161571	AL591936 Homo sapiens chromo	gb_htg:AC026940	+	38.00	71.49	1.0e+05	185399	AC026940	Homo sapiens chr
gb_pr:CN051R1G	-	38.00	72.70	8.9e+04	161813	AL163194 Human chromosome 14	gb_htg:AP001887	-	38.00	71.49	1.0e+05	185405	AP001887	Homo sapiens chr
gb_in:AC008356	-	38.00	72.65	8.9e+04	162595	AC008356 Drosophila melanog	gb_htg:AC084273	-	38.00	71.47	1.0e+05	18		

gb\_htg:AC079273 + 38.00 71.45 1.0e+05 186296 ! AC079273 Mus musculus chromo  
gb\_htg:AC023830 + 38.00 71.42 1.0e+05 186846 ! AC023830 Homo sapiens chromo  
gb\_pr:AC090824 - 38.00 71.41 1.0e+05 187160 ! AC090824 Homo sapiens chromo  
gb\_htg:AL359982 + 38.00 71.38 1.0e+05 187691 ! AL359982 Homo sapiens chromo  
gb\_htg:AC025999 + 38.00 71.36 1.0e+05 188116 ! AC025999 Homo sapiens chromo  
gb\_htg:AC092711 + 38.00 71.36 1.0e+05 188216 ! AC092711 Mus musculus clone  
- 38.00 71.36 1.0e+05 188219 ! AC022875 Homo sapiens clone  
gb\_htg:AC022875 - 38.00 71.35 1.0e+05 188287 ! AL023693 Human DNA sequence  
gb\_pr:HS32B1 + 38.00 71.35 1.0e+05 188491 ! AP003084 Homo sapiens chromo  
em\_htg\_hum:AC027725 + 38.00 71.31 1.1e+05 189153 ! AC027725 Homo sapiens clone  
gb\_htg:AP001333 + 38.00 71.31 1.1e+05 189181 ! AP001333 Homo sapiens chromo

seq\_name: gb\_pat:AX005896

seq\_documentation\_block:  
LOCUS AX005896 51 bp DNA 24-AUG-2000  
DEFINITION Sequence 105 from Patent WO9909174.  
ACCESSION AX005896  
VERSION AX005896.1 GI:9928885  
KEYWORDS  
SOURCE bread wheat.  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

REFERENCE  
AUTHORS Harberd,N.P. and Peng,J.  
TITLE Genetic control of plant growth and development  
JOURNAL Patent: WO 9909174-A 105 25-FEB-1999;

FEATURES  
source Location/Qualifiers  
1..51  
/organism="Triticum aestivum"  
/db\_xref="taxon:4565" 7 t

BASE COUNT 7 a 16 c 21 g 7 t

ORIGIN  
|||||  
1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
273 GACGAGCTGTCGGCGCTCGGGTACAGGTGCGCGCTCCGACATGGC 322

alignment\_scores:  
Quality: 81.00 Length: 17  
Ratio: 4.765 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-485-529-104 x AX005896 ..  
Align seg 1/1 to: AX005896 from: 1 to: 51

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
1 GACGAGCTGTCGGCGCTCGGGTACAGGTGCGCGCTCCGACATGGC 50

17 a 17

51 G 51

seq\_name: gb\_pat:AX005857

seq\_documentation\_block:  
LOCUS AX005857 453 bp DNA 24-AUG-2000  
DEFINITION Sequence 66 from Patent WO9909174.  
ACCESSION AX005857  
VERSION AX005857.1 GI:9928852  
KEYWORDS  
SOURCE bread wheat.  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

REFERENCE  
AUTHORS Harberd,N.P. and Peng,J.  
TITLE Genetic control of plant growth and development  
JOURNAL Patent: WO 9909174-A 66 25-FEB-1999;

FEATURES  
source Location/Qualifiers  
1..453  
/organism="Triticum aestivum"  
/db\_xref="taxon:4565"

misc\_feature 6 /note="n is any nucleotide"

misc\_feature 11 /note="n is any nucleotide"

misc\_feature 46 /note="n is any nucleotide"

misc\_feature 85 /note="n is any nucleotide"

misc\_feature 107 /note="n is any nucleotide"

misc\_feature 445 /note="n is any nucleotide"

misc\_feature 449 /note="n is any nucleotide"

BASE COUNT 85 a 136 c 158 g 67 t 7 others  
ORIGIN

alignment\_scores:  
Quality: 81.00 Length: 17  
Ratio: 4.765 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-485-529-104 x AX005857 ..

Align seg 1/1 to: AX005857 from: 1 to: 453

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
273 GACGAGCTGTCGGCGCTCGGGTACAGGTGCGCGCTCCGACATGGC 322

17 a 17

323 G 323

seq\_name: gb\_pl:TAE242531

seq\_documentation\_block:  
LOCUS TAE242531 1872 bp DNA PLN 28-JUL-1999  
DEFINITION Triticum aestivum rht-Dla gene for gibberellin response modulator.  
ACCESSION AJ242531  
VERSION AJ242531.1 GI:5640156  
KEYWORDS gibberellin response modulator; rht-Dla gene.  
SOURCE bread wheat.  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 1872)  
Peng,J., Richards,D.E., Hartley,N.M., Murphy,G.P., Devos,K.M.,  
Flintham,J.E., Beales,J., Fish,L.J., Worland,A.J., Pelica,F.,  
Sudhakar,D., Christou,P., Snape,J.W., Gale,M.D. and Harberd,N.P.  
'Green revolution' genes encode mutant gibberellin response  
modulators  
Nature 400 (6741), 256-261 (1999)  
99347734  
2 (bases 1 to 1872)  
Harberd,N.P., Peng,J. and Richards,D.E.  
Green revolution genes encode mutant gibberellin response  
modulators  
Unpublished  
3 (bases 1 to 1872)  
Richards,D.E.  
Direct Submission  
Submitted (25-MAY-1999) Richards D.E., Molecular Genetics, John  
Innes Centre, Colney Lane, NR4 7UJ, UNITED KINGDOM  
Location/Qualifiers

seq\_name: gb\_pat:AX005857

seq\_documentation\_block:  
LOCUS AX005857 453 bp DNA 24-AUG-2000  
DEFINITION Sequence 66 from Patent WO9909174.  
ACCESSION AX005857  
VERSION AX005857.1 GI:9928852  
KEYWORDS  
SOURCE bread wheat.  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 453)  
Harberd,N.P. and Peng,J.  
Genetic control of plant growth and development  
Patent: WO 9909174-A 66 25-FEB-1999;

```
source
1..1872
/organism="Triticum aestivum"
/db_xref="taxon:4565"
1..1872
/genes="rht-D1a"
1..1872
/genes="rht-D1a"
/genes="GAI ortholog"
/codon_start=1
/product="gibberellin response modulator"
/protein_id="CAB51555.1"
/db_xref="GI:5640157"
/translation="MKREYQDAGSGGGGSGGSKMMVSAAGEEVDELLAAL
GYKVRASMDADVAKLEAMGMGGVAGAAPDDSFATHLATDTVHYNPTDLSSWV
ESMLSELNAPPPLPPALNASTSVTSGGYEDLPPSYDSSSIYALRPISPAG
ATAPADLSADSVRDKRMTGSGTSSSSSSSLGGGARSVEAAPPPAAANATP
ALPVVVDQTQAGIRLVHALLACAEAVQDENLSAEALVKQIPLLAASQGGAMRKVAA
YFGALARRVFRFPQDSSLLDAFAADLLHAHFYESCPLYLKFHFHTANQAILFAFAG
CRRVHVDGFKQGMQWPAALQALALRPGPSFRLTGVGPPQDPTDALQQVWKLAA
QFAHTIRVDFOYRGLVAATLADLPFMLQPEGEDPNEPEVIAVNSVFEMHRLAOP
GALEKVLGTVRAVRPRIVTVVEQANHNSTFELDRFTESLHYSTMEDSLEGGSSGG
PSEVSSGAAAPAAAGIDQVMSEVYLGROIKNVACGAERTERHETLGQWRNLGNA
GFETVHLGNSNAYKQASTLLALFAGDGGYKVEEGCLTLGLWHTRLPLIATSARWLAP"
BASE COUNT 265 a 700 c 641 g 266 t
ORIGIN

alignment_scores:
Quality: 81.00 Length: 17
Ratio: 4.765 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-485-529-104 x TAE242531 ..
Align seg 1/1 to: TAE242531 from: 1 to: 1872
1 AspCluleuLeuAlaLaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
112 GACGAGCTGCTGGCGGCTCGGCTACAGGTGGCGCCCTCCGACATGGC 161
17 a 17
162 G 162
seq_name: gb_pat:AX005805
seq_documentation_block:
2125 bp DNA PAT 24-AUG-2000
LOCUS AX005805
DEFINITION Sequence 14 from Patent WO9909174.
ACCESSION AX005805
VERSION AX005805.1 GI:9928802
KEYWORDS
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 2125)
AUTHORS
Harberd,N.P. and Peng,J.
TITLE
Genetic control of plant growth and development
JOURNAL
Patent: WO 9909174-A 14 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
Location/Qualifiers
source
1..2125
/organism="Triticum aestivum"
/db_xref="taxon:4565"
misc_feature
6 /note="n is any nucleotide"
misc_feature
11 /note="n is any nucleotide"
misc_feature
46 /note="n is any nucleotide"
misc_feature
85 /note="n is any nucleotide"
misc_feature
107 /note="n is any nucleotide"
misc_feature
407 /note="n is any nucleotide"
misc_feature
417 /note="n is any nucleotide"
misc_feature
430 /note="n is any nucleotide"
misc_feature
436 /note="n is any nucleotide"
misc_feature
448 /note="n is any nucleotide"
misc_feature
452 /note="n is any nucleotide"
misc_feature
490 /note="n is any nucleotide"
misc_feature
498 /note="n is any nucleotide"
misc_feature
556 /note="n is any nucleotide"
misc_feature
562 /note="n is any nucleotide"
misc_feature
727 /note="n is any nucleotide"
misc_feature
769 /note="n is any nucleotide"
misc_feature
1444 /note="n is any nucleotide"

BASE COUNT 325 a 768 c 723 g 309 t
ORIGIN

alignment_scores:
Quality: 81.00 Length: 17
```

```
Ratio: 4.765 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-485-529-104 x AX005805 ..
Align seg 1/1 to: AX005805 from: 1 to: 2125
1 AspCluleuLeuAlaLaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
139 GACGAGCTGCTGGCGGCTCGGCTACAGGTGGCGCCCTCCGACATGGC 188
17 a 17
189 G 189
seq_name: gb_pat:AX005794
seq_documentation_block:
2709 bp DNA PAT 24-AUG-2000
LOCUS AX005794
DEFINITION Sequence 3 from Patent WO9909174.
ACCESSION AX005794
VERSION AX005794.1 GI:9928799
KEYWORDS
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 2709)
AUTHORS
Harberd,N.P. and Peng,J.
TITLE
Genetic control of plant growth and development
JOURNAL
Patent: WO 9909174-A 3 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
Location/Qualifiers
source
1..2709
/organism="Triticum aestivum"
/db_xref="taxon:4565"
misc_feature
6 /note="n is any nucleotide"
misc_feature
11 /note="n is any nucleotide"
misc_feature
46 /note="n is any nucleotide"
misc_feature
85 /note="n is any nucleotide"
misc_feature
107 /note="n is any nucleotide"
misc_feature
407 /note="n is any nucleotide"
misc_feature
417 /note="n is any nucleotide"
misc_feature
430 /note="n is any nucleotide"
misc_feature
436 /note="n is any nucleotide"
misc_feature
448 /note="n is any nucleotide"
misc_feature
452 /note="n is any nucleotide"
misc_feature
490 /note="n is any nucleotide"
misc_feature
498 /note="n is any nucleotide"
misc_feature
556 /note="n is any nucleotide"
misc_feature
562 /note="n is any nucleotide"
misc_feature
727 /note="n is any nucleotide"
misc_feature
769 /note="n is any nucleotide"
misc_feature
1444 /note="n is any nucleotide"
```



```
misc_feature      /note="n is any nucleotide"
1450
/note="n is any nucleotide"
1556
/note="n is any nucleotide"
1742
/note="n is any nucleotide"
1808
/note="n is any nucleotide"
1900
/note="n is any nucleotide"
1940
/note="n is any nucleotide"
1982
/note="n is any nucleotide"
2438
/note="n is any nucleotide"
2604
/note="n is any nucleotide"
2638
/note="n is any nucleotide"
2681
/note="n is any nucleotide"
2689
/note="n is any nucleotide"
2706
/note="n is any nucleotide"
31 others
BASE COUNT      473 a 900 c 823 g 482 t
ORIGIN

alignment_scores:
  Quality: 81.00      Length: 17
  Ratio: 4.765       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-485-529-104 x AX005794 ..
Align seg 1/1 to: AX005794 from: 1 to: 2709

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
273 GACGAGCTGCTGGCGCGCTCGGTACAAGTGCCTCGCATGGC 322

17 a 17
323 G 323

seq_name: gb_pat:AX005807

seq_documentation_block:
LOCUS AX005807 302 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 16 from Patent WO9909174.
ACCESSION AX005807
VERSION AX005807.1 GI:9928804
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 302)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 16 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source
 1..302
  /organism="Zea mays"
  /db_xref="taxon:4577"
BASE COUNT      52 a 82 c 123 g 45 t
ORIGIN
```

```
alignment_scores:
  Quality: 78.00      Length: 17
  Ratio: 4.588       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-485-529-104 x AX005807 ..
Align seg 1/1 to: AX005807 from: 1 to: 302

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
100 GATGAGCTGCTGGCGCGCTCGGTACAAGTGCCTCGCATGGC 149

17 a 17
150 G 150

seq_name: gb_pat:AX005808

seq_documentation_block:
LOCUS AX005808 371 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 17 from Patent WO9909174.
ACCESSION AX005808
VERSION AX005808.1 GI:9928805
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 371)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 17 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source
 1..371
  /organism="Zea mays"
  /db_xref="taxon:4577"
BASE COUNT      56 a 121 c 139 g 55 t
ORIGIN

alignment_scores:
  Quality: 78.00      Length: 17
  Ratio: 4.588       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-485-529-104 x AX005808 ..
Align seg 1/1 to: AX005808 from: 1 to: 371

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
64 GATGAGCTGCTGGCGCGCTCGGTACAAGTGCCTCGCATGGC 113

17 a 17
114 G 114

seq_name: gb_pat:AX005803

seq_documentation_block:
LOCUS AX005803 770 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 12 from Patent WO9909174.
ACCESSION AX005803
VERSION AX005803.1 GI:9928800
KEYWORDS
SOURCE
  Oryza sativa.
```

ORGANISM *Oryza sativa*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; *Oryza*.  
REFERENCE 1 (bases 1 to 770)  
AUTHORS Harberd,N.P. and Peng,J.  
TITLE Genetic control of plant growth and development  
JOURNAL Patent: WO 9909174-A 12 25-FEB-1999;  
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)  
FEATURES  
source 1..770  
          /organism="Oryza sativa"  
          /db\_xref="taxon:4530"  
BASE COUNT 104 a 252 c 294 g 120 t  
ORIGIN

alignment\_scores:  
Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:  
US-09-485-529-104 x AX005803 ..  
Align seg 1/1 to: AX005803 from: 1 to: 770

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
114 GACGAGCTGCTGGGGCGCTCGGGTACAGGTCGGTCTCGCAGATGGC 163

17 a 17  
|  
164 c 164

seq\_name: gb\_pl.ZMA242530

seq\_documentation\_block:  
LOCUS ZMA242530 1890 bp DNA PLN 28-JUL-1999  
DEFINITION *Oryza sativa* partial d8 gene for gibberellin response modulator.  
ACCESSION AJ242530  
VERSION AJ242530.1 GI:5640154  
KEYWORDS D8 gene; gibberellin response modulator.  
SOURCE *Oryza sativa*.  
ORGANISM *Oryza sativa*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
ciade; Panicoideae; Andropogoneae; *Oryza*.  
REFERENCE 1 (bases 1 to 1890)  
AUTHORS Peng,J., Richards,D.E., Hartley,N.M., Murphy,G.P., Devos,K.M.,  
Flintham,J.E., Beales,J., Fish,L.J., Worland,A.J., Pellica,F.,  
Sudhakar,D., Christou,P., Snape,J.W., Gale,M.D. and Harberd,N.P.  
TITLE 'Green revolution' genes encode mutant gibberellin response  
modulators  
JOURNAL Nature 400 (6741), 256-261 (1999)  
MEDLINE 99347734  
REFERENCE 2 (bases 1 to 1890)  
AUTHORS Richards,D.E.  
TITLE Direct Submission  
JOURNAL Submitted (25-MAY-1999) Richards D.E., Molecular Genetics, John  
Innes Centre, Colney Lane, NR4 7UJ, UNITED KINGDOM  
FEATURES  
source 1..1890  
          /organism="Oryza sativa"  
          /db\_xref="taxon:4577"  
          1..1890  
          /feature="d8"  
          1..>1890  
          /feature="d8"  
          /note="GAI ortholog"  
          /codon\_start=1  
          /product="gibberellin response modulator"  
          /protein\_id="CAB51557.1"

308 GATGAGCTGCTGCGCGCTCGGTACAAGGTGCGTTCGTCGGATATGCG 357

17 a 17

358 G 358

seq\_name: gb\_pl:AB030956

seq\_documentation\_block: 2500 bp mRNA PLN 30-MAR-2000  
LOCUS AB030956  
DEFINITION Oryza sativa mRNA for OsGAI, complete cds.  
ACCESSION AB030956  
VERSION AB030956.1 GI:6970471  
KEYWORDS OsGAI.  
SOURCE Oryza sativa (cultivar:Nipponbare) seedling cDNA to mRNA.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (sites)  
Ogawa,M., Kusano,T., Katsumi,M. and Sano,H.  
Rice gibberellin-insensitive gene homolog, OsGAI, encodes a  
nuclear-localized protein capable of gene activation at  
transcriptional level  
Gene 245 (1), 21-29 (2000)  
JOURNAL MEDLINE  
20179680  
REFERENCE 2 (bases 1 to 2500)  
Sano,H., Ogawa,M. and Kusano,T.  
Direct Submission  
Submitted (09-AUG-1999) to the DDBJ/EMBL/GenBank databases.  
Mikihiro Ogawa, Nara Institute of Science and Technology, Plant  
Molecular Breeding; Takayama 8916-5, Ikoma, Nara 630-0101, Japan  
(E-mail:m-ogawa@nibs.aist-nara.ac.jp, Tel:+81-743-72-5652,  
Fax:+81-743-72-5659)

FEATURES

source  
1..2500  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/dev\_stage="seedling"  
216..2093  
/gene="OsGAI"  
216..2093  
/gene="OsGAI"  
/note="GAI homolog"  
/codon\_start=1  
/product="OsGAI"  
/protein\_id="BAA90749.1"  
/db\_xref="GI:6970472"  
/translation="MKREYQAGSGSGGSDMGSCDKVMAGAAGEEDVDDELLAA  
LGKYVRSDMADVAQKLEOLEMAMGVSAPGAADGCFVSHLATDTHVYNPSLSSW  
VESMLSENLPLPIPPAPPAARHASTSTVTGGSGGFELPAADSSSTYALRPI  
SLPVVATADPSAADSARDTKMRITGGGSSSSSSSSSLGGASRGSVVEAPPATQC  
AAANAAPVVPVVDTEAGIRLVHALLACAEVQOENFAAAELVKOIPTLAASQGG  
AMRVVAYFGALARRVFRPADSTLLDAFADLLHAHFESCPYLKFAHTANQAI  
LEAFAGRHVVDFGKQGMQWPAQLALRPGGPPSFLRTGVPQPDTEALQO  
VWKLQAFHTRIVDFQYGLVAAATLADLEPFMLQPEGEADANEPEVIAVNSVFELH  
RLLAQGALEKVLCTHVAREPRITVVQENHNHSGSFLDFETSLHYSTFMDPSLEG  
GSSQGAELSPAAAGGGGTDQVMSEVILGRQICNVVACEGNERETRIETLIGQWRNLG  
RAGFEPVHLGNSATYKQASTLLALFAGGQYRVEKEGCLTLGWHTRLIATSAWRVAA  
A"

BASE COUNT 413 a 861 c 771 g 455 t

ORIGIN

alignment\_scores:

Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x AB030956 ..

FEATURES

Align seg 1/1 to: AB030956 from: 1 to: 2500

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
330 GACGAGCTGCTGCGCGCTCGGTACAAGGTGCGTTCGTCGCACATGCG 379

17 a 17

380 C 380

seq\_name: gb\_pl:AC087797

seq\_documentation\_block:  
LOCUS AC087797 122497 bp DNA PLN 03-MAY-2001  
DEFINITION Oryza sativa chromosome 3 BAC OSJNB0022E02 genomic sequence,  
complete sequence.  
ACCESSION AC087797  
VERSION AC087797.5 GI:13699786  
KEYWORDS HTG.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 122497)  
Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N.,  
Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M.,  
Tsitrin,T., Riggs,F., Hsiao,J., Zismann,V., Blunt,S., Pal,G.,  
Vanaken,S.E., Utterback,T.R., Feldblyum,T.V., Quackenbush,J.,  
Salzberg,S.L., White,O. and Fraser,C.M.  
Oryza sativa chromosome 3 BAC OSJNB0022E02 genomic sequence  
Unpublished  
2 (bases 1 to 122497)  
Buell,R.  
Direct Submission  
Submitted (25-JAN-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
3 (bases 1 to 122497)  
Buell,R.  
Direct Submission  
Submitted (20-APR-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
4 (bases 1 to 122497)  
Buell,R.  
Direct Submission  
Submitted (03-MAY-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuella@tigr.org  
On Apr 20, 2001 this sequence version replaced gi:13605985.  
Address all correspondence to:rice@tigr.org  
BAC clone OSJNB0022E02 is from Oryza sativa chromosome 3  
The orientation of the sequence is from SP6 to T7 end of the BAC  
clone.  
Genes were identified by a combination of several methods: Gene  
prediction programs including Fgenesh (<http://www.softberry.com/>),  
genscan and Genscan+ (Chris Burge,  
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,  
<http://genemark.biology.gatech.edu/genemark/>), and Gensplice  
(Mihaila Perlea and Steven Salzberg, [contact.mperlea@tigr.org](http://contact.mperlea@tigr.org)),  
searches of the complete sequence against a peptide database and  
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as unknown proteins.  
Genes without protein or EST similarity, that are predicted by more  
than two gene prediction programs over most of their length are  
annotated as hypothetical proteins. Genes encoding tRNAs are  
predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are  
identified by RepeatMasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).  
Location/Qualifiers

```

source
1. 122497
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="3"
/map="C944"
/clone="OSJNB0022E02"
/complement(<1271..>1579)
/gene="OSJNB0022E02.11"
/complement(1271..1579)
/notes="OSJNB0022E02.11"
/notes="predicted by fgenesh"
/complement(1271..1579)
/gene="OSJNB0022E02.11"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50135.1"
/db_xref="GI:13937304"
/translation="MNTKAAPLVSSPFSVLSPAADVPRPTLRAPPLHPLGLSVAP
SCRTRSSGTLWTPVKTALMAKVVVDDGWSHVSFAQKWAFLAENLAGSVAANTS
"
<1386..>1640
/gene="OSJNB0022E02.18"
1386..1640
/notes="OSJNB0022E02.18"
/notes="predicted by glimmerR"
1386..1640
/gene="OSJNB0022E02.18"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50133.1"
/db_xref="GI:13937302"
/translation="MRASVETGVHNVPLERVQLGATLRPGCGSGARNVGRGSTAA
GSDNTENGEDTNGAFAVRSITAPAPQNTGTGATVREEL"
complement(join(<5240..5517,5605..5684,6132..6212,
6314..6383,6539..6613,6693..6898,7117..7426,7506..7551,
8773..8842,9564..9633,11257..11341,11440..11524,
11608..11698,12255..>12557))
/gene="OSJNB0022E02.2"
/complement(5240..12557)
/gene="OSJNB0022E02.2"
/notes="EST D47656, AU088605 from this gene"
complement(join(5439..5517,5605..5684,6132..6212,
6314..6383,6539..6613,6693..6898,7117..7426,7506..7551,
8773..8842,9564..9633,11257..11341,11440..11524,
11608..11698,12255..12388))
/gene="OSJNB0022E02.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAK50136.1"
/db_xref="GI:13937305"
/translation="MORRRQAQTWAGVKTAQAAAAHAALFCFTLLIALKVDGRTAYSW
WIIFLWLFHGIIVARFSMPALPHGRHWAPCHSIVAAPLLIAFAELLCIYLES
RVKSVTDVTLKVFLLPAFEVILADNFRMCALMPGDESMDEAIWETLPHFWVA
ISNVFLATTTLLKLDGVDGALGWDLFINTGIAECFAFLVCTWFNPMHKSNP
GEASSSAIRVEDWESGLLPSLEDEHEQERLCGLDIGHVNMKIPLVIFQVLLCMRL
EATPPAQIPIPALFSPFILOAGVLFSLARLEKVLRLNRNGPSPNYLTFISKV
RDCFALRHQIGSEITKCTKQEFERLQNEKVLRCRYEGEICMVLLPCRHRTLCKTC
SDKKCKPCICRVPFIEEMPVYDV"
complement(12328..12420)
/product="hypothetical protein"
/rpt_family="CGGGGG"
/rpt_family="GA"
join(<14168..15023,15131..15274,15344..16687,16818..17803,
17887..18048,18262..>18366)
/gene="OSJNB0022E02.12"
14168..18366
/notes="OSJNB0022E02.12"
/notes="similar to gag pol polyprotein GB:AAD12997
GI:4234852 (Zea mays)"
join(14168..15023,15131..15274,15344..16687,16818..17803,
17887..18048,18262..>18366)

repeat_region
repeat_region
mRNA
gene
CDS

17887..18048,18262..18366)
/gene="OSJNB0022E02.12"
/codon_start=1
/product="putative gag-pol polyprotein"
/protein_id="AAK50117.1"
/db_xref="GI:13937286"
/translation="MAGFADALRPDKFTGVHFKRWQIRVTLMLTAMKCFWVSTGKPEG
VLTAEQOKQFEATTLFVGCILSVGLDRIVEVYMHMTDAKELMDALNTKFGATDASND
FYTALQHKRQEVSEGLIASLDEVERKAREKDAASKGDSOSSANVHQAQNSKQYK
ACTTTFKQKKNNPNQDERTCFVCGQVGHARLAKCPQPKGMAKAGOTSANVTI
GNTDGSYGRTGFRHPNGEVIWTCFCSCWCHGRSEVIFGKRAAEERACPFTRQESC
YEGGLFRSLDFCNKSNHICGSDDEANVHSHRCHINFGIMSLRSLMCLIPKFS
IVKSKCHVQSQKQPKRPHKAAERNLAPLELLHSDLCMEMGVLTKGKRKYFTMLID
DATRFVYLLTKDALDYFKIYKAEVENOLDKIKRLSRSDRGGEFFSNFDFCEE
HGIHERTPYSPESNGIAERKNRTITDLVNAMLDTAGLPKAWMGALLTSHVLNRV
PNRKDKTPYEIWIHGRKPSLSYLRTWGLCAKVNVPITKKRKLGPKTVDCLGYAHS
IAYRLIVSEEDVSEAPRRSKRQRTAKSGDDFTVYLVDDTPKSISEASPADYWK
TEVHSEMDSTIANGTWERTPYGCKPVGCKWFKKRLRPDASHGLLVHOMDYKTAFL
NGELDEIYMDQDPGVFVEQEGCKLLKSLYGLQAQKQMHKEKFDKTTISAGFAVN
EADKCVYRHGGEGVTLQYVDDILIFGNLEVINEVKSFSONDKDLGVADVIL
NIKLIRGENTILLOSHYVEKILNRECYIDSKPSPTPYDPSLILRLKRNKRNQLEYS
QIIGSLMYLASATRPDISFAVSKLSRFTSNPGDDHRRALERVMRYLKGTVELGLHTG
YPALLEGYSDSNWISDVDEIKATSGYVFTLGGGAVSWRSCQITILTRSTMEAEITLD
TATVAEWLRLDLMLPVVEKPTCEKTIIEVCQEIKKILRSYNVGLHPNSEKPGRLHEG
TITKCDRQCIEGDFETHILEGTLCELDWCWSQSMKILGESSRKLTLDLGV"
complement(18822..18858)
/rpt_family="(CAGGC)n"
complement(19612..19640)
/rpt_family="AT_rich"
complement(20459..20487)
/rpt_family="AT_rich"
<21271..>21576
/gene="OSJNB0022E02.13"
21271..21576
/gene="OSJNB0022E02.13"
/notes="predicted by fgenesh"
21271..21576
/gene="OSJNB0022E02.13"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50120.1"
/db_xref="GI:13937289"
/translation="MTTRTVLLSEYSDSHDNGRNGAVAGVVGHVGGVAGGRLHGYI
ACAVLLASTPYAPLHQDELOHQVMTAKDCRWQAPAWDVDEYETSNDASVPNS"
complement(22507..22529)
/rpt_family="(TA)n"
complement(22544..22588)
/rpt_family="AT_rich"
complement(23124..23217)
/rpt_family="(GGAGAA)n"
complement(join(<23165..23588,23765..23917,25379..>25551))
/gene="OSJNB0022E02.14"
complement(23165..25551)
/gene="OSJNB0022E02.14"
/notes="predicted by fgenesh"
complement(join(23165..23588,23765..23917,25379..25551))
/gene="OSJNB0022E02.14"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50123.1"
/db_xref="GI:13937292"
/translation="MAAGSGARAVGNGGGVGGGPPAPASLPPLPRHPSPRRRTPA
SPGGLLPTRLRCQYTGSLPESAAGHAYATPRSPRRPGADADEGEVLGLSSAAAV
WMGAQARSCEGNAESMAVALSACPYAGDALAKGATGCGVKGKVIHGYVFTITSL
VCVYKLGEMDNNAKAFRADATEKNIVTWNTLITSYATAGLCEALDVLQAEQIGGTV"

alignment_scores:
Quality: 78.00 Length: 17
Ratio: 4.588 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

```

```
alignment_block:
US-09-485-529-104 x AC087797 ..
Align seg 1/1 to: AC087797 from: 1 to: 12497

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
|||||
69197 GAGCAGCTGTCGGCGCGTCGGGTACAGGTGCGGTCCGACATGGC 69246
17 a 17
1
69247 C 69247

seq_name: gb_pat:AX005810

seq_documentation_block: 725 bp DNA PAT 24-AUG-2000
LOCUS AX005810
DEFINITION Sequence 19 from Patent WO9909174.
ACCESSION AX005810
VERSION AX005810.1 GI:9928807
KEYWORDS
SOURCE
ORGANISM
Oryza sativa.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 725)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 19 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source
1..725
/organism="Oryza sativa"
/db_xref="taxon:4530"
misc_feature 171
/note="n is any nucleotide"
misc_feature 302
/note="n is any nucleotide"
misc_feature 427
/note="n is any nucleotide"
misc_feature 444
/note="n is any nucleotide"
misc_feature 459
/note="n is any nucleotide"
misc_feature 711
/note="n is any nucleotide"
BASE COUNT 98 a 226 c 276 g 119 t 6 others
ORIGIN

alignment_scores:
Quality: 72.00 Length: 16
Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 93.750

alignment_block:
US-09-485-529-104 x AX005810 ..
Align seg 1/1 to: AX005810 from: 1 to: 725

2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
|||||
109 GAGCTGTCGGCGCGTCGGGTACAGGTGCGGTCCGACATGGCC 156

seq_name: gb_pat:AX005867

seq_documentation_block:
LOCUS AX005867 324 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 76 from Patent WO9909174.
ACCESSION AX005867
VERSION AX005867.1 GI:9928862
KEYWORDS

SOURCE
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 324)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 76 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source
1..324
/organism="Triticum aestivum"
/db_xref="taxon:4565"
misc_feature 158
/note="n is any nucleotide"
misc_feature 161
/note="n is any nucleotide"
misc_feature 217
/note="n is any nucleotide"
BASE COUNT 39 a 141 c 85 g 56 t 3 others
ORIGIN

alignment_scores:
Quality: 70.00 Length: 15
Ratio: 4.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-485-529-104 x AX005867/rev ..
Align seg 1/1 to reverse of: AX005867 from: 1 to: 324

3 LeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
|||||
156 CTGCTGGCGCGCTCGGTACAGGTGCGCGCTCCGACATGGCG 112

seq_name: gb_pl:ATRG2

seq_documentation_block: 1951 bp mRNA PLN 19-AUG-1997
LOCUS ATRG2
DEFINITION A.thaliana mRNA for RGA2 gene.
ACCESSION Y11337
VERSION Y11337.1 GI:2339979
KEYWORDS RGA2 gene.
SOURCE thale.cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1951)
AUTHORS Truong,H.N., Caboche,M. and Daniel-Vedele,F.
TITLE Sequence and characterization of two Arabidopsis thaliana cDNAs
isolated by functional complementation of a yeast gln3 gdh1 mutant
JOURNAL FEBS Lett. 410 (2-3), 213-218 (1997)
MEDLINE 97379310
REFERENCE
2 (bases 1 to 1951)
AUTHORS Truong,H.N.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1997) H.N. Truong, INRA-Versailles, Laboratoire
de Biologie Cellulaire, Route de Saint-Cyr, 78026 Versailles Cedex,
FRANCE
FEATURES
source
1..1951
/organism="Arabidopsis thaliana"
/sub_species="L). Heynh"
/db_xref="taxon:3702"
/dev_stage="two-leaf stage seedlings"
/note="Landsberg erecta ecotype"
196..1794
/gene="RGA2"
196..1794
CDS
```



```
/gene="RGA2"  
/codon_start=1  
/product="RGA2 protein"  
/protein_id="CAA72178.1"  
/db_xref="GI:2339980"  
/db_xref="SPTREMBL:O23643"  
/translation="MKRDHHHHQDKKTKMMNEEDGNGMDELLAVLGVKVRSEMA  
VAQKLEQVMSNVQEDDLSQLATETVHYNPAELTYWLDLSMLTDLPSSNAEYDLK  
AIPGDAILNQFAIDSASSNQGGGDTYTNKRLKCSNGVETTTATAESTRHVVLVD  
SOENGVRLVHALLACAEAVOKENLTVAEALVKQIGFLAVSOIGAMROVATYFAEALR  
RIVRLSPSPIDHSLSDTLQMHFYETCPYLKFAHFTANOAILLEAFQGKKRVHVIDFS  
MSQGLQWPAALMQALRPGPPVFRITGIGPPAPDNFDYHLHEVGCKLAHLAEAHVFEF  
EYRGFVANTLADLASMELRPESEISVAVNSVFEHLKILGRPGAIDKVLGVVNOIKP  
EFTVVEQSNHNSPIFLDRETESLHYSTLDSLEGVPSGQDKVMSEVYLGKQICNV  
VACDGPDRVERHETLSQWRNRFGSAGFAAAHIGSNAFKQASMLLALFNGGEGYRVEES  
DGCLMLGWHTRPLIATSAWKLSTN"
```

BASE COUNT 482 a 419 c 475 g 575 t  
ORIGIN

alignment\_scores:  
Quality: 70.00 Length: 17  
Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x ATRGA2 ..

Align seg 1/1 to: ATRGA2 from: 1 to: 1951

```
1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
274 GATGAGCTTCTAGCTGTTCTTGGTTACAGGTTAGTTCATCGGAATGCG 323
```

17 a 17

324 T 324

seq\_name: gb\_pat:A64697

seq\_documentation\_block: 1964 bp DNA PAT 29-MAR-1999  
LOCUS A64697  
DEFINITION Sequence 1 from Patent WO9729123.  
ACCESSION A64697  
VERSION A64697.1 GI:4530762

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1964)

Harberd,N.P., Peng,J., Carol,P. and Richards,D.E.

NUCLEIC ACID ENCODING GAI GENE OF ARABIDOPSIS THALIANA

Patent: WO 9729123-A 1 14-AUG-1997;

INNES JOHN CENTRE INNOV LTD (GB)

Other publication AU 1799697 19970828.

COMMENT Location/Qualifiers

FEATURES

source  
1..1964  
/organism="Arabidopsis thaliana"  
/strain="LANDSBERG ERECTA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/map="20.9"  
/haplotype="GAI"

BASE COUNT 489 a 426 c 474 g 575 t  
ORIGIN

alignment\_scores:  
Quality: 70.00 Length: 17  
Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x A64697 ..

Align seg 1/1 to: A64697 from: 1 to: 1964

```
1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
287 GATGAGCTTCTAGCTGTTCTTGGTTACAGGTTAGTTCATCGGAATGCG 336
```

17 a 17

337 T 337

seq\_name: gb\_pl:ATY15193

seq\_documentation\_block: 1964 bp DNA PLN 01-NOV-1997  
LOCUS ATY15193  
DEFINITION Arabidopsis thaliana GAI gene.  
ACCESSION Y15193  
VERSION Y15193.1 GI:2569937

KEYWORDS GAI gene.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1964)

Peng,J., Carol,P., Richards,D.E., King,K.E., Cowling,R.J.,

Murphy,G.P. and Harberd,N.P.

The Arabidopsis GAI gene defines a signalling pathway that

negatively regulates gibberellin responses

Genes Dev. In press

2 (bases 1 to 1964)

Harberd,N.P.

Direct Submission

Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular

Genetics, Colney Lane, Norwich, NR4 7UJ, UK

Location/Qualifiers

source

1..1964  
/organism="Arabidopsis thaliana"

/strain="Landsberg erecta"

/db\_xref="taxon:3702"

/chromosome="1"

209..1807  
/gene="GAI"

209..1807  
/gene="GAI"

/codon\_start=1

/protein\_id="CAA75492.1"

/db\_xref="GI:2569938"

/db\_xref="SPTREMBL:O23724"

/translation="MKRDHHHHQDKKTKMMNEEDGNGMDELLAVLGVKVRSEMA  
VAQKLEQVMSNVQEDDLSQLATETVHYNPAELTYWLDLSMLTDLPSSNAEYDLK  
AIPGDAILNQFAIDSASSNQGGGDTYTNKRLKCSNGVETTTATAESTRHVVLVD  
SOENGVRLVHALLACAEAVOKENLTVAEALVKQIGFLAVSOIGAMRKVATYFAEALR  
RIVRLSPSPIDHSLSDTLQMHFYETCPYLKFAHFTANOAILLEAFQGKKRVHVIDFS  
MSQGLQWPAALMQALRPGPPVFRITGIGPPAPDNFDYHLHEVGCKLAHLAEAHVFEF  
EYRGFVANTLADLASMELRPESEISVAVNSVFEHLKILGRPGAIDKVLGVVNOIKP  
EFTVVEQSNHNSPIFLDRETESLHYSTLDSLEGVPSGQDKVMSEVYLGKQICNV  
VACDGPDRVERHETLSQWRNRFGSAGFAAAHIGSNAFKQASMLLALFNGGEGYRVEES  
DGCLMLGWHTRPLIATSAWKLSTN"

BASE COUNT 489 a 426 c 474 g 575 t  
ORIGIN

alignment\_scores:  
Quality: 70.00 Length: 17  
Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:  
US-09-485-529-104 x ATY15193 ..

Align seg 1/1 to: ATY15193 from: 1 to: 1964

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 287 GATGAGCTCTAGCTGTTGGTTACAAAGTTAGGTCATCGGAATGGC 336

17 a 17  
 337 T 337

seq\_name: gb\_pl.ATY15194

seq\_documentation\_block: 2201 bp mRNA PLN 01-NOV-1997  
 LOCUS ATY15194  
 DEFINITION Arabidopsis thaliana mRNA for GRS protein.  
 ACCESSION Y15194  
 VERSION Y15194.1 GI:2569939  
 KEYWORDS GRS protein.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 2201)  
 AUTHORS Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J.,  
 Murphy, G.P., and Harberd, N.P.

TITLE The Arabidopsis GAI gene defines a signalling pathway that  
 negatively regulates gibberellin responses

JOURNAL Genes dev. In press  
 REFERENCE 2 (bases 1 to 2201)  
 AUTHORS Harberd, N.P.

TITLE Direct Submission  
 JOURNAL Submitted (22-OCT-1997) N.P. Harberd, John Innes Centre, Molecular  
 Genetics, Colney Lane, Norwich, NR4 7UJ, UK

FEATURES  
 source

1..2201  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="2"  
 144..1907

/codon\_start=1  
 /product="GRS protein"  
 /protein\_id="CAA75493.1"  
 /db\_xref="GI:2569940"  
 /db\_xref="SPTREMBL:O23725"

CDS

/translation="MKRDHGFQGRLSNHGTSSTSSSSSKDKMMVKKEEDGGGNMDD  
 ELLAVGKVRSEMAEVALKLEQETMMSNVQEDGLSHLATDTHVNPSELYSWLND  
 MLELNPPPLPSSNGLDPLVLPSPICGFPASDYDLVPGNAIQFPAIDSSSSNN  
 QNRLKSCSPDSMVTSTGTQIGGVIGTATTTTAAESTRVSAILVDSQENG  
 VRLVHALMACAEAIQNNLTAEALVKQICGLAVSQAGAMRKVATYFAEALARIYRL  
 SPQNIQDCLSDTLQMHFYETCPYLKFAHFAHTANOAILAEAFEGKRVHVIDFSMNQGL  
 QMPALMDALREGGPPFTRLTGIGPPAPDNDHLHEYGCKLAQLAEAHVFEFFYRGF  
 VANSIADLASMLELRPSDTEAVAVNSVFELHKLGRPGGIEKVLGVVVKQIKPVIFTV  
 VEQSNHNGPVLDRFTESLHYSTLFDLSLEGVPNSODKVMSEVYLGKQICNLVACEG  
 PDRVERHETLSQWGNRFGSSGLAPAHLSNAFAKQASMLLSVFNSSQGYRVEESNGCLM  
 LGWHTRPLITTSNAKLSRAAY"  
 BASE COUNT 576 a 450 c 525 g 650 t  
 ORIGIN

alignment\_scores:  
 Quality: 70.00 Length: 17  
 Ratio: 4.118 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x ATY15194

Align seg 1/1 to: ATY15194 from: 1 to: 2201

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17

|||||  
 273 GATGAGCTCTCGCTGTTAGGTACAAAGTTAGGTCATCGGAGATGGC 322

17 a 17  
 323 G 323

seq\_name: gb\_pl.ATRGAL

seq\_documentation\_block: 2210 bp mRNA PLN 19-AUG-1997  
 LOCUS ATRGAL  
 DEFINITION A.thaliana mRNA for RGAL gene.  
 ACCESSION Y11336  
 VERSION Y11336.1 GI:2339977  
 KEYWORDS RGAL gene.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 2210)  
 AUTHORS Truong, H.N., Caboche, M. and Daniel-Vedele, F.

TITLE Sequence and characterization of two Arabidopsis thaliana cDNAs  
 isolated by functional complementation of a yeast gln3 gdh1 mutant  
 JOURNAL FEBS Lett. 410 (2-3), 213-218 (1997)  
 MEDLINE 97379310

REFERENCE 2 (bases 1 to 2210)  
 AUTHORS Truong, H.N.

TITLE Direct Submission

JOURNAL Submitted (13-FEB-1997) H.N. Truong, INRA-Versailles, Laboratoire  
 de Biologie Cellulaire, Route de Saint-Cyr, 78026 Versailles Cedex,  
 FRANCE

FEATURES  
 source

1..2210  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /sub\_species="(L). Heynh"  
 /db\_xref="taxon:3702"  
 /dev\_stage="two-leaf stage seedlings"  
 /note="Landsberg erecta ecotype"  
 132..1895

gene

CDS

/gene="RGAL"  
 132..1895  
 /gene="RGAL"  
 /codon\_start=1  
 /product="RGAL protein"  
 /protein\_id="CAA72177.1"  
 /db\_xref="GI:2339978"

/translation="SPTREMBL:O23642"  
 /translation="MKRDHGFQGRLSNHGTSSTSSSSSKDKMMVKKEEDGGGNMDD  
 ELLAVGKVRSEMAEVALKLEQETMMSNVQEDGLSHLATDTHVNPSELYSWLND  
 MLELNPPPLPSSNGLDPLVLPSPICGFPASDYDLVPGNAIQFPAIDSSSSNN  
 QNRLKSCSPDSMVTSTGTQIGGVIGTATTTTAAESTRVSAILVDSQENG  
 VRLVHALMACAEAIQNNLTAEALVKQICGLAVSQAGAMRKVATYFAEALARIYRL  
 SPQNIQDCLSDTLQMHFYETCPYLKFAHFAHTANOAILAEAFEGKRVHVIDFSMNQGL  
 QMPALMDALREGGPPFTRLTGIGPPAPDNDHLHEYGCKLAQLAEAHVFEFFYRGF  
 VANSIADLASMLELRPSDTEAVAVNSVFELHKLGRPGGIEKVLGVVVKQIKPVIFTV  
 VEQSNHNGPVLDRFTESLHYSTLFDLSLEGVPNSODKVMSEVYLGKQICNLVACEG  
 PDRVERHETLSQWGNRFGSSGLAPAHLSNAFAKQASMLLSVFNSSQGYRVEESNGCLM  
 LGWHTRPLITTSNAKLSRAAH"  
 BASE COUNT 583 a 454 c 518 g 655 t  
 ORIGIN

alignment\_scores:  
 Quality: 70.00 Length: 17  
 Ratio: 4.118 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x ATRGAL

Align seg 1/1 to: ATRGAL from: 1 to: 2210

```
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
261 GACGAGCTTCTCGCTGTTTGGTTACAAAGTTAGTTCATCGGAGATGCG 310

17 a 17
|
311 G 311

seq_name: gb_pl:ATA224957

seq_documentation_block:
LOCUS   ATA224957      4081 bp      DNA           PLN           19-MAY-1998
DEFINITION   Arabidopsis thaliana RGAL gene.
ACCESSION   AJ224957
VERSION     AJ224957.1 GI:3021335
KEYWORDS    RGAL gene.
SOURCE      thale cress.
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 4081)
Sanchez-Fernandez,R., Ardiles-Diaz,W., Van Montagu,M., Inze,D. and
May,M.J.
Cloning of the Arabidopsis thaliana RGA-like gene, a putative
member of the VHIID domain transcription factor family
Unpublished
2 (bases 1 to 4081)
Sanchez-Fernandez,R.
Direct Submission
Submitted (14-MAR-1998) Sanchez-Fernandez R., Department of
Genetics, University of Gent (VIB), K.L. Ledeganckstraat 35, 9000
Gent, BELGIUM

FEATURES             Location/Qualifiers
     source           1..4081
                     /organism="Arabidopsis thaliana"
                     /cultivar="Col-0"
                     /db_xref="taxon:3702"
     gene             join(900..1048,1130..1343,1434..1529,1881..3410)
                     /gene="RGA-like"
     gene             join(900..1048,1049..1129,1130..1343,1344..1433,
                     1434..1529,1530..1880,1881..3410)
                     /gene="SCR2"
                     900..1048
                     /gene="SCR2"
                     /number=1
     CDS              join(900..1048,1130..1343,1434..1529,1881..3410)
                     /gene="RGA-like"
                     /codon_start=1
                     /protein_id="CAA12242.1"
                     /db_xref="GI:3021336"
                     /db_xref="SPREMBL:O65367"
                     /translation="MEEVSEMEVEVQNQLDSSPAQNVKFKGLKNSIQTFNGSDYV
                     FQIVKIDTAIVSLSTNTKLYSPVTQGYGECKGSDTVNQIAFSSDASAAPHVL
                     HSCSGDSTMTQVSRIDTNDQDIEFISYGGADNLLAGCKEQRHNRHRE
                     SSAEGGSSMTTVIKEEAGVDELLVLGIVKVRSSMDADVAKHLEQMLVLDGILSN
                     LSDETHYNPDSLGVESMLSDLPTRIQEKPDSEYDLRAIPCSAVYPERDVTBRS
                     KREIESELSTSVSVLDSQETGVRVLHALLACAAEQVQNNKLADALVKKHVGLLAS
                     SQAGMKWTATYFAEGLARIYRIYPRDDVASSFSDTLQIHFEYSCPYLKKFAHFTAN
                     QALLEVFAETAEKRVHVDLGLNGLQWLPALIQALRNPGPPDFRLTIGYSLTDIQEV
                     GWKLGOLATIGVNFEEKSLNLLNLDLKPMLDIRPGLSEVAVNSVFLHRLIAPHG
                     SIDKFLSTKISRPDITWVVEQANHNTVFLDRFTESLHYSSLSFDSLEGGPPSDQRY
                     MSELFLGROIILNLVACEGDEVRVERHETLNQNRNRFGLGGFKPYSIGSNAYKQASMLLA
                     LYAGADGVNVEEGCLLGLWQTRPLIATSAWRINRV"
                     1049..1129
                     /gene="SCR2"
                     /number=1
     intron           1130..1343
     exon             /gene="SCR2"
                     /number=2
     intron           1344..1433
                     /gene="SCR2"
                     /number=2
```

```
1434..1529
/gene="SCR2"
/number=3
1530..1880
/gene="SCR2"
/number=3
1881..3410
/gene="SCR2"
/number=4

BASE COUNT      1152 a      752 c      862 g      1315 t
ORIGIN

alignment_scores:
  Quality:   70.00      Length:   17
  Ratio:     4.118      Gaps:     0
Percent Similarity: 100.000 Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x ATA224957 ..
Align seg 1/1 to: ATA224957 from: 1 to: 4081

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
1968 GACGAGCTTCTCGCTGTTTGGTTACAAAGTTAGTTCATCGGAGATGCG 2017

17 a 17
|
2018 T 2018

seq_name: gb_pl:ATK3M16

seq_documentation_block:
LOCUS   ATK3M16      26604 bp      DNA           PLN           07-AUG-2000
DEFINITION   Arabidopsis thaliana DNA chromosome 5, BAC clone K3M16 (ESSA
project).
ACCESSION   AL391150
VERSION     AL391150.1 GI:9755767
KEYWORDS    thale cress.
SOURCE      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 24604)
Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asanizu,E., Kotani,H.,
Tabata,S., Mewes,H.W., Rudd,S., Lemcke,K. and Mayer,K.F.X.
Unpublished
2 (bases 1 to 24604)
EU Arabidopsis sequencing,project.
Direct Submission
Submitted (07-AUG-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
Location/Qualifiers
     source           1..26604
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
     gene             90..3656
                     /gene="K3M16_10"
     exon            90..236
                     /gene="K3M16_10"
                     /number=1
     CDS              join(90..236,706..858,954..1017,1270..1349,1432..1579,
                     1936..1997,2237..2299,2675..2733,3037..3116,3204..3247,
                     3335..3559,3638..3656)
                     /gene="K3M16_10"
```

```
/note="similarity to cisplatin resistance-associated
overexpressed protein - Homo sapiens, EMBL:AB034205"
/codon_start=1
/product="putative protein"
/protein_id="CAC01888.1"
/db_xref="GI:9755768"
/translation="MDAMRKOLDVLMGANRGDYTEVNRKYDRDVCRLYLSGLCPHE
LFQLTAKGVNDYDEDAIDRLVECDRIKGRALNRLQEDAKAAIAISVTEFTQSP
EILELSQIKEMKEADLHDLGKTDLRIKALELVEEMRTKRADLQAVLLIDAFNKDR
TSLQOAVPAQPPAATLPPDPRTQEMINEKLKAEELGEOGMVDEAKALEAEALKK
LTAQEPVVDSTKYTAADVRIQDKLRLCDICGAFLSVYDSRDLADHFGKHLGYM
LIRKLAELQEKKNYKHKEVERESRSRERSRSGDGGDNRDRGRDVRDRSRDR
DRHHDHREHNRNYNQSRYDSRSRSRSRSRERPRDHRRRRRHRY"
/237..705
intron
/gene="K3M16_10"
/number=1
exon
706..858
/gene="K3M16_10"
/number=2
intron
859..953
/gene="K3M16_10"
/number=2
exon
954..1017
/gene="K3M16_10"
/number=3
intron
1018..1269
/gene="K3M16_10"
/number=3
exon
1270..1349
/gene="K3M16_10"
/number=4
intron
1350..1431
/gene="K3M16_10"
/number=4
exon
1432..1579
/gene="K3M16_10"
/number=5
intron
1580..1935
/gene="K3M16_10"
/number=5
exon
1936..1997
/gene="K3M16_10"
/number=6
intron
1998..2236
/gene="K3M16_10"
/number=6
exon
2237..2299
/gene="K3M16_10"
/number=7
intron
2300..2674
/gene="K3M16_10"
/number=7
exon
2675..2733
/gene="K3M16_10"
/number=8
intron
2734..3036
/gene="K3M16_10"
/number=8
exon
3037..3118
/gene="K3M16_10"
/number=9
intron
3119..3203
/gene="K3M16_10"
/number=9
exon
3204..3247
/gene="K3M16_10"
/number=10
intron
3248..3334
/gene="K3M16_10"
/number=10
exon
3335..3359
/gene="K3M16_10"
/number=11
exon
3360..3637
/gene="K3M16_10"
/number=11
exon
3638..3656
/gene="K3M16_10"
/number=12
exon
3657..5704
/gene="K3M16_20"
/number=1
gene
5627..6984
/gene="K3M16_20"
/number=2
CDS
join(5627..5704,6047..6122,6689..6984)
/note="strong similarity to farnesylated protein ATPF6 -
Arabidopsis thaliana, EMBL:U64909
Contains Prenyl group binding site (CAAX box) AAL46-149"
/codon_start=1
/product="farnesylated protein ATPF6-like protein"
/protein_id="CAC01889.1"
/db_xref="GI:9755769"
/translation="MGAFDYISSCSYTYANAKTKRKPLQTVDIKVMKMDGCGERRVR
NVVRRMGKVSVEVNRKQSRITVNGHVDPNKVLKRVKSTGKKAEEFWYIPOHMYVYFP
APGMYDKRAPAGHIRNPTQSFPTANAPEENVVSLFSDDNVHAACIN"
/5705..6046
/gene="K3M16_20"
/number=1
exon
6047..6122
/gene="K3M16_20"
/number=2
intron
6123..6688
/gene="K3M16_20"
/number=2
exon
6689..6984
/gene="K3M16_20"
/number=3
gene
7823..9866
/gene="K3M16_30"
complement(join(7823..8600,8690..8795,8926..8974,
9062..9142,9681..9866))
/gene="K3M16_30"
complement(join(7823..8600,8690..8795,8926..8974,
9062..9142,9681..9866))
/gene="K3M16_30"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC01890.1"
/db_xref="GI:9755770"
/translation="MVFPACVQCGRONPCRKVVGPTLGFVAPLITGIIIEWPVGAVV
YIFKHAKRRIRIMHPARKDSIKKAVKREKDVRAKTLISILYRNQKITKSLAFVSDNFC
SKLVQVVDQVFEVMPQRYVLFMTLLKFSFFGIEYTLVGRGKDYNLFGNKPQDDDFRK
TWKEMEDDDTLWSGSESDNDKTKQESGRNRLEIKKARQAAKSTQDQIDGDSDB
LYSVWSGSDREKTLMTGEGDDDDIPEPRPNEASDKYLDKLFEPFEKPKYRTISEL
LKSENEPEELSPPGKKARKLAVENALKKLNKGPDGRYTNVWEVNSDVLILIGAFENIIS
GPYEYELRKGGPKRLNQMFQFDKIOTRMRDPNFKFTPEIKLKPSKLVPRKKWQAQSR
IRKAKR"
complement(7823..8600)
/gene="K3M16_30"
/number=1
intron
complement(8601..8689)
/number=1
exon
complement(8690..8795)
/gene="K3M16_30"
/number=2
intron
complement(8796..8925)
/number=2
exon
complement(8926..8974)
/gene="K3M16_30"
/number=3
intron
complement(8975..9061)
/number=3
exon
complement(9062..9142)
/gene="K3M16_30"
/number=4
```

```

intron      complement(9143..9680)
            /number=4
exon        complement(9681..9866)
            /gene="K3M16_30"
            /number=5
exon        complement(11207..11647)
            /gene="K3M16_40"
            /number=1
gene        complement(11207..11647)
            /gene="K3M16_40"
gene        11207..11647
            /gene="K3M16_40"
            /gene="K3M16_40"
            /note="strong similarity to calmodulin-related protein"

alignment_scores:
  Quality: 70.00      Length: 17
  Ratio: 4.118      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x ATK3M16/rev ..
Align seg 1/1 to reverse of: ATK3M16 from: 1 to: 26604
1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
16029 GACGAGTTCCTGCTGGTTGGTTACAAAGTTCGATCTTCAGACATGCG 15980
17 a 17
15979 A 15979

```

```
seq_name: gb_pl:ATAC009895
```

```

seq_documentation_block:
LOCUS      ATAC009895 82289 bp      DNA            PLN            24-JAN-2001
DEFINITION Arabidopsis thaliana chromosome III BAC T21P5 genomic sequence,
            complete sequence.
ACCESSION  AC009895
VERSION    AC009895.4 GI:12408719
KEYWORDS   HTG.
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsi.
            Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.,
            Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
            Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
            Arabidopsis thaliana chromosome III BAC T21P5 genomic sequence
            Unpublished
            Lin,X. and Kaul,S.
            2 (bases 1 to 82289)
            Direct Submission
            Submitted (04-SEP-1999) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
            3 (bases 1 to 82289)
            Lin,X.
            Direct Submission
            Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
            On Jan 24, 2001 this sequence version replaced gi:12280866.
            Address all correspondence to:
            Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr.
            Rockville, MD 20850, USA
            e-mail: xlin@tigr.org
            BAC clone T21P5 is from Arabidopsis chromosome III and is near the
            molecular marker ml172.

```

```

FEATURES             source
1..82289
     /organism="Arabidopsis thaliana"
     /cultivar="Columbia"
     /db_xref="taxon:3702"
     /chromosome="III"
     /map="ml172"
     /clone="T21P5"
1..24217
     /note="overlap with BAC clone T12J13
     (AC009327:65718..89934)."
     complement(join(<8..215,451..587,675..857,959..1050,
1459..1593,1682..1748,1837..2055,2146..2340,2492..2879,
2979..>3100))
     /gene="T21P5.1"
     complement(<8..>3100)
     /gene="T21P5.1"
     /note="similar to hypothetical protein GB:CAB38918
     [Arabidopsis thaliana]"
     complement(join(8..215,451..587,675..857,959..1050,
1459..1593,1682..1748,1837..2055,2146..2340,2492..2879,
2979..3100))
     /gene="T21P5.1"
     /note="unknown protein"
     /codon_start=1
     /protein_id="AAF01600.1"
     /db_xref="GI:6017117"
     /translation="MSQIISGNNNIPLSEVYVSLVKNKDKFKIRDLPFVRSRYEN
YFKYFVYTLWKFKQENKQLVEAGLKEWGEIGEIASRIAQLYGYHMYTSDAGYLS
LIDVEYFALITRFKDLFDOLNANKLRLFLARLWCLVLGRWVHOLVDQFKR
LIDCKRTFQETDFKENVAAQVEIRFLKSDTAFMNIIRYSLVLPNDLADGPPFR
RSRLTDAILSYCYNEVKYSELTDFSRMLQCLEWEPFSGLSYQSTGAKMGQNPVGV
ARINQSQMDPTLPPNKAVLYRPSITHEFLAVLTICEELPSHGILLLYLSAGSKIG
QISSPILSARSATSVENILRDSEPTIKOETEPSLOITPSGSSYIYPSDLVPTTRKP
LIIFIDSDSTVFNKICGAKGEPAALLSPSHPTPLISADFSRQPSGSLVPTLTPSP
VQAFCLLSEISADMETDIFTAEKLSSSNWASLATSDTLHPVMSQLDKDFLLR
RLLRFTFCRAVLALYTPVFNKNQNPQCECPSPLESLPTAPAVQSAVFQMANVFGAT
SKFTIPQDITMLESF"
     join(<3627..3661,3762..3843,3918..3995,4081..4140,
4240..4344,4437..4547,4629..4670,4860..4973,5097..5254,
5361..5494,5576..5682,5785..5827,5903..>6144)
     /gene="T21P5.2"
     <3627..>6144
     /gene="T21P5.2"
     /note="predicted by genscan"
     join(3627..3661,3762..3843,3918..3995,4081..4140,
4240..4344,4437..4547,4629..4670,4860..4973,5097..5254,
5361..5494,5576..5682,5785..5827,5903..6144)
     /gene="T21P5.2"
     /note="hypothetical protein"

```

The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of three methods: Gene Prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://www.cbs.stanford.edu/~chris/GENSCANW.html>), and NetPlantGene (<http://www.cbs.stanford.edu/~netpene/cbsnetgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tldb/at/at.html>).

Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are identified by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).

Regions of the genome identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

#### FEATURES

```

Location/Qualifiers
1..82289
     /organism="Arabidopsis thaliana"
     /cultivar="Columbia"
     /db_xref="taxon:3702"
     /chromosome="III"
     /map="ml172"
     /clone="T21P5"
1..24217
     /note="overlap with BAC clone T12J13
     (AC009327:65718..89934)."
     complement(join(<8..215,451..587,675..857,959..1050,
1459..1593,1682..1748,1837..2055,2146..2340,2492..2879,
2979..>3100))
     /gene="T21P5.1"
     complement(<8..>3100)
     /gene="T21P5.1"
     /note="similar to hypothetical protein GB:CAB38918
     [Arabidopsis thaliana]"
     complement(join(8..215,451..587,675..857,959..1050,
1459..1593,1682..1748,1837..2055,2146..2340,2492..2879,
2979..3100))
     /gene="T21P5.1"
     /note="unknown protein"
     /codon_start=1
     /protein_id="AAF01600.1"
     /db_xref="GI:6017117"
     /translation="MSQIISGNNNIPLSEVYVSLVKNKDKFKIRDLPFVRSRYEN
YFKYFVYTLWKFKQENKQLVEAGLKEWGEIGEIASRIAQLYGYHMYTSDAGYLS
LIDVEYFALITRFKDLFDOLNANKLRLFLARLWCLVLGRWVHOLVDQFKR
LIDCKRTFQETDFKENVAAQVEIRFLKSDTAFMNIIRYSLVLPNDLADGPPFR
RSRLTDAILSYCYNEVKYSELTDFSRMLQCLEWEPFSGLSYQSTGAKMGQNPVGV
ARINQSQMDPTLPPNKAVLYRPSITHEFLAVLTICEELPSHGILLLYLSAGSKIG
QISSPILSARSATSVENILRDSEPTIKOETEPSLOITPSGSSYIYPSDLVPTTRKP
LIIFIDSDSTVFNKICGAKGEPAALLSPSHPTPLISADFSRQPSGSLVPTLTPSP
VQAFCLLSEISADMETDIFTAEKLSSSNWASLATSDTLHPVMSQLDKDFLLR
RLLRFTFCRAVLALYTPVFNKNQNPQCECPSPLESLPTAPAVQSAVFQMANVFGAT
SKFTIPQDITMLESF"
     join(<3627..3661,3762..3843,3918..3995,4081..4140,
4240..4344,4437..4547,4629..4670,4860..4973,5097..5254,
5361..5494,5576..5682,5785..5827,5903..>6144)
     /gene="T21P5.2"
     <3627..>6144
     /gene="T21P5.2"
     /note="predicted by genscan"
     join(3627..3661,3762..3843,3918..3995,4081..4140,
4240..4344,4437..4547,4629..4670,4860..4973,5097..5254,
5361..5494,5576..5682,5785..5827,5903..6144)
     /gene="T21P5.2"
     /note="hypothetical protein"

```





Genes were identified by a combination of several methods: Gene prediction programs including Genscan+ (Chris Burge, <http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky, <http://genemark.biology.gatech.edu/GeneMark/>), GlimmerA (a variant of GlimmerX, see Mihalea Perteza, <http://www.tigr.org/sofclab/glimmerm.htm>), and GeneSplicer (Mihalea Perteza and Steven Salzberg, contact [mperteza@tigr.org](mailto:mperteza@tigr.org)), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

Location/Qualifiers

1. 85702

source

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="1"

/clone="T27F4"

complement(join(<1757..1964,2044..2152,2244..2346,2431..2465,2573..2790,2869..2984,3105..3357,3447..3534,3625..3702,3870..3945,4033..4088,4206..4275,4478..4642))

/gene="T27F4.2"

complement(1757..4642)

/gene="T27F4.2"

/note="similar to GB:AAB64244 from [Arabidopsis thaliana] (Plant Mol. Biol. 34 (1), 57-68 (1997))"

complement(join(1757..1964,2044..2152,2244..2346,2431..2465,2573..2790,2869..2984,3105..3357,3447..3534,3625..3702,3870..3945,4033..4088,4206..4275,4478..4642))

/codon\_start=1

/product="beta-glucosidase, putative; 4642-1757"

/protein\_id="AAG52157.1"

/db\_xref="GI:12324390"

/translation="MALOKFLPLMLGILLTILSVTTAVDDPVCATSKLSRASPPNGFLFGPTAAFOVEGAINETCGPALWDLYCENPERCSGDHADVADEPHRYKEDIQLMKNLNDFAFRLSIANSRIFPHRGKEGVSQAGYQYHELIDELKNGIVPFTVFEHWDTPQDLEDEGGSLSDNIFREYADYVFTYGGKVKNNWITENPWFVFAHAGYDLGKKAPGRCRSYVPGCEDREGSGKEAYLVSHNLLNAHAEVFEVFRQYKGGKIGIAHSPAWFEPHDLKQSDNAPTYSRVLDLFWLHEPTTSGDYQIQMKDLGLYRLPQFTAAQAKLVKSTDFVGLNYVTSTFSNYPKDPSPKSPWKODSLVSNPKNVDSHAIGSMPLTAALPVYAKGFRLLKYIKDYANPEITMENGYGDKLGTDSVDVCGTADHNKRYILQRLHLLAMNEAICDKRVITGTFVWSLLDNFEWQDGIKRNFGLIYVDFKNNLTRYEKESAKYIKDFLAQGVRSALRKDEL"

repeat\_region

complement(6634..7675)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6978..7293)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6981..7112)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6983..7139)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6983..7166)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6983..7085)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6983..7115)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(6991..7205)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7001..7085)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7002..7088)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7036..7205)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7038..7208)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7055..7139)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7065..7202)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7082..7166)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7090..7205)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7092..7169)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7109..7231)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7119..7200)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7136..7328)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7199..7276)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7200..7284)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7201..7358)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7237..7356)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7281..7355)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(7286..7355)

/rpt\_family="SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1|SIMPLEGUY1 is a putative non-autonomous DNA transposon - a consensus."

complement(79126,9231..9448,9526..9641,9735..9987,10076..10163,10253..10330,10454..10529,10859..10914,10992..11061,11220..>11384))

/gene="T27F4.3"

complement(8406..11384)

/gene="T27F4.3"

/note="similar to GB:AAB64244 from [Arabidopsis thaliana] (Plant Mol. Biol. 34 (1), 57-68 (1997))"

complement(join(8406..8613,8690..8798,8882..8984,9092..9126,9231..9448,9526..9641,9735..9987,10076..10163,10253..10330,10454..10529,10859..10914,10992..11061,11220..>11384))

mRNA

gene

complement(8406..11384)

/note="similar to GB:AAB64244 from [Arabidopsis thaliana] (Plant Mol. Biol. 34 (1), 57-68 (1997))"

complement(join(8406..8613,8690..8798,8882..8984,9092..9126,9231..9448,9526..9641,9735..9987,10076..10163,10253..10330,10454..10529,10859..10914,10992..11061,11220..>11384))

CDS

```
11220..11384))
/gene="T27F4.3"
/codon_start=1
/product="beta-glucosidase, putative; 11384-8406"
/protein_id="AA52159.1"
/db_xref="GI:12324392"
/translat="MALQKPLDGLLFLITIVVSSITAVDPCPTTSKLSRASFPNG
FVFCTATAAFQVEGAINETCRGPAIMDFCKRNPERCSGHNADVAVDFPHRYKEDIQL
MKNLNTDAPLSLTAWSRIPIPHRGKGVSOAGVFYHDLIDELKNGIIPFVVFHWD
TPQDLDEYGFSLSENIVKDFEYADYVFTEYGGKKNWITFNEPWFVAHAGYDVCKK
APGRCSRYLKCCEDRGSRSEYALYSHNLLNHAHAENEVFRQKYGGKIGIAHSPAW
FEPHDUKSDNDVPTVSKVLDFMLGWLHDPITFDGYFOIMKDLGLHRLPFRFTSOKALP
KSTDFVGLNYITSTFNHNKPPDPTSPKQDSLVAWEKPNVDYHSAIGSQPLTAALP
VYAKGFRSLKYIKDYANPEIMMENGYGDKLKDSDSEVGADYNRKYYLQRLHLLA
MNEAICIDKRVTVGYFVWSLLDNFEMQDGYNNRFGLYYVDFKNNLTRYEKESAKYKD
FLGGVPSALKKDEL"
complement(join(<14953..15216,15304..15444,15534..>16490))
/gene="T27F4.4"
complement(14953..16490)
/gene="T27F4.4"
/note="contains Pfam profile: PF00646 F-box domain"

alignment_scores:
  Quality: 70.00      Length: 17
  Ratio: 4.118      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x AC020665 ..
Align seg 1/1 to: AC020665 from: 1 to: 85702
1 AspGluLeuAlaLeuGlyTyrIysValArgAlaserAspMetal 17
|||||
50067 GACGACGTTTGGTTGTTTAGTTACAAAGTTCGATCATCCGACATGCC 50116

17 a 17
50117 T 50117

seq_name: gb_pl:AC005560
seq_documentation_block:
LOCUS AC005560 95137 bp DNA PLN 05-APR-2000
DEFINITION Arabidopsis thaliana chromosome II section 3 of 255 of the complete
sequence. Sequence from clones F10A8, F219.
ACCESSION AC005560 AE002093
VERSION AC005560.2 GI:5598472
KEYWORDS HTG.
SOURCE
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 95137)
Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
Fuji,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
Feldblyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Renning,C.M.,
Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Unayam,L.,
Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
Venter,J.C.
Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana
Nature 402 (5763), 761-768 (1999)
JOURNAL MEDLINE 20083487
PUBMED 10617197
REFERENCE 2 (bases 1 to 95137)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
```

## COMMENT

On Dec 17, 1999 this sequence version replaced gi:3785968. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/tdb/at/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<ftp://arthur.epm.ornl.gov/pub/xgrail>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://gnomic.stanford.edu/GENSCANW.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/WashU/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [at@tigr.org](mailto:at@tigr.org).

FEATURES	Location/Qualifiers
source	1..95137 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="II"
misc_feature	<1..300 /note="Sequence from clone F10A8"
misc_feature	301..>95137 /note="Sequence from clone F219"
mRNA	<341..>997 /gene="At2g01390"
gene	<341..>997 /gene="At2g01390"
CDS	/note="F219.1; predicted by genscan" 341..997 /gene="At2g01390" /note="hypothetical protein" /codon_start=1 /protein_id="AAC67316.1" /db_xref="GI:3785968"
repeat_region	complement(1085..1127) /rpt_family="(TAAAA)n" join(<1175..1294,1620..>1940) /gene="At2g01400"
mRNA	<1175..>1940 /gene="At2g01400"
gene	/note="F219.2; predicted by genscan"

SLGTYDVLSSLEKTSDFQKEVALLRKKSLVASARFENVHVEDKICDLLPATNLL  
\*  
LAFDYSLEMGTHLKSLAYLALIGNFLSNELPKVIEYKEMVKAHSLGCGYQGAHLIH  
RLGFRRLAADVDFLLPDDQKGAAYATLMDVYISAGSPKAMKTLRMRREIMP  
\*  
complement(1085..1127)  
/rpt\_family="(TAAAA)n"  
join(<1175..1294,1620..>1940)  
/gene="At2g01400"  
/note="F219.2; predicted by genscan"

```
CDS
join(1175..1294,1620..1940)
/genes="At2g01400"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAC67317.1"
/db_xref="GI:3785970"
/translation="MARSCSIWTPVLISLSPVTGSGPKISRRRVILATISGSPQLLE
AKIKPQKQVLDSKDRNTMLYLAAGFLGGINFLNGEAAEARVGRKENRKALKLEK
LRAKASEPNKNSGNQKTEKELEKEEVPPLPPPLVVEANLIQ"
complement(<2069..>3232)
/genes="At2g01410"
complement(<2069..>3232)
/genes="At2g01410"
complement(<2069..>3232)
/note="F219.3; predicted by genefinder"
complement(2069..3232)
/genes="At2g01410"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAC67318.1"
/db_xref="GI:3785971"
/translation="MKANNSSSSTRYRSFFFPILFAVLFLGLLIPSSADNRHVINP
RSPGLYPEGLTWDPRDHPFLVGLSHSRTTHSVSDAGVIEITLISDLDLPENSTILGLAV
DSTNRLLACIOSLPPLPFPFALASVDLRSGRRVFLSPSLPGDEDIARDVANDV
AYDFKGNATVTSANKNFVKVDKDGASIFSKSPFNQSPQADADADASRDGCLNGIV
YISGVLVYQSGTKVFKVDEDSGNARLVLLNGDLIAADGMTRRRRDGTVMVYSOKK
LWLAKSQDSWSGCVYDEIDLDIEGFPPTAVTAVGRDRIYVLXGRVMEGMSGYKKEG
AREWFGIEBVSEKEGEDKIMLYVLIGFGFAVFCFWRFOMKKLIINMDKKIT"
complement(join(6967..7404,7509..7585,7680..7837,
8215..8300,8440..8707,8865..>10059))
/genes="At2g01420"
complement(6967..>10059)
/genes="At2g01420"
/note="F219.4"
complement(join(7338..7404,7509..7585,7680..7837,
8215..8300,8440..8707,8865..>10059))
/genes="At2g01420"
/codon_start=1
/product="putative auxin transport protein"
/protein_id="AAC67319.1"
/db_xref="GI:3785972"
/translation="MITWHDLTVLTAVPLVLYAMILAVSGVQWVKIFSPDQCSGILNR
FVAIFAPVLSSHFSTNDPYAMNFRFAADTLQRIIMVLALWANLTKNGSLEWMI
TIFSLTPTNLTWPLPLIAMYGTVAGSLMVQVVVLCIIWYTLFLFVYRGAKLL
INEQFPTCAISVRFKGVSDVSLDGHDFLEDAEIGNDKLHVTVRKNSASRLMM
TPRPSNLTAETVLSSTPRSGNFHSDYFVYSGVPGGRLSNFGPADLYSVQSSRGPT
PRPSPEENNAKYGFYNNTSSVPAGSYPAFPNPFSTGTVGSKPKNKPENQOQL
QKDSKSHADKELHFWSSASPVSDVFGGGAGNVATEEQSEQAKETKIPNIVSDQP
RKSNAFGGDDIGLSDSGEGEREIEKATAGLNKMGNSNFAELEAAGDGGGNVTHMP
PTSVTRLLIMVWRKLRINPNYISLIGLIWALVAYRWHPAKTIQOSISILSDAG
LGMWFSGLFNLALQPKIACGNSVATFAMAVRFTGPAIMAVAGIAIGLHGLLRIA
IVQAALPQGIVPFVFAKEYNVHPTILSTGVIFGMLIALPITLVYIILLG"
14152..14264
/rpt_family="POLY_A"
complement(join(<14658..14916,16258..16337,16458..>16607))
/genes="At2g01430"
complement(<14658..>16607)
/genes="At2g01430"
/note="F219.5"
complement(join(14658..14916,16258..16337,16458..>16607))
/genes="At2g01430"
/codon_start=1
/product="putative homeodomain transcription factor"
/protein_id="AAC67320.1"
/db_xref="GI:3785973"
/translation="MNRLPSSDEDDDEESHDDGSPAPPKKLLKLTREQSRLLEDSPRQ
NHTLPNKQEVIAKHLMLPRQIEVFNQNRARRSKLKQTEMECEYLKRWFSLTEENH
RLHREVEELRAMKVGPVTNVSASLTMCPCRERTVPAASPSRAVVPVPAKTFPPQER
DR"
complement(join(<20954..21559,21680..22027,
22234..22377,22526..22606,22742..22813,23017..23250,
23352..23441,23559..23640,23906..24069,24308..24542,
24746..24858,24942..>25415))
/genes="At2g01440"
complement(<20954..>25415)
/genes="At2g01440"
/note="F219.6"
complement(join(20954..21559,21680..22027,
22234..22377,22526..22606,22742..22813,23017..23250,
23352..23441,23559..23640,23906..24069,24308..24542,
24746..24858,24942..>25415))
/genes="At2g01440"
/codon_start=1
/product="putative ATP-dependent DNA helicase (RECG)"
/protein_id="AAC67321.1"
/db_xref="GI:3785974"
/translation="MECDNVDDFIDKKSDQVKKDLVLACKRFPFIIILDSRPRVELYS
NKSYSSESSILKATPTDNSFLPTPMHGWFDPDLNLSRTLSFCELIQNDSDSPRED
ILDDGSFTSKTATSEVATSDVFAAQRELATSIDMPLGSKRHSNQLDSCGFHTMK
KLLHFPRTVADIQNAQVDIEDGXYLIFVYGVKLVSSKGVRRASSFSFLEIVTSCVSGR
DRTPEDLSHNAEDKAGKSFILHLKKFPGTRFTWQPFNLNIOEKHKVGDVLVCISGVK
SLAEDHFEEMREYNIDVLDEESSHRAQGRPYIPYSGKLPKFLDVISRALRVL
PANMDPIPKETIKVFLGSLNDAYVGIHEPTLDEADLARKRLIFDEFFYLQARLYO
MLQSLGKTEKDVLLEKFRKPLVINSYIEEWSLTITKSLKALPYSLTSPQSLSAVSEI
WDLKRPVNMRLQGVGGKTVAFVACMEVIGSGVQAFAFMAPTELLAHOYEOCRD
LLENMEDLSQSAISFTLIGTHSLIAEKIEVSALRIAVDQQRFGVQRCFKNSKLYG
TSMISKSGSDSDTTSKADLSMAPHVLAMSATIPRSLALALYGDLSLTITGMPIGR
IPVETHIFEGNETGIKEVYSMMLEDLKSGRVYVYVPIVDQSQLPOLRAASAELEIV
```

```
alignment_scores
Quality: 70.00 Length: 17
Ratio: 4.118 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353

alignment_block
US-09-485-529-104 x AC005560/rev ..

Align seg 1/1 to reverse of: AC005560 from: 1 to: 95137

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
84075 GACGAGCTCTCGCTGTTTAGTTACAAAGTTAGTCATCGGAGATGCG 84026

17 a 17
|
84025 G 84025

seq_name: gb_pl:AC006917

seq_documentation_block:
LOCUS AC006917 132999 bp DNA PLN 28-JUN-2000
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
I, complete sequence.
ACCESSION AC006917
VERSION AC006917.6 GI:4757662
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 132699)
Chao,Q., Shinn,P., Dunn,P., Buehler,E., Kahn,S., Kim,C., Walker,M.,
Williams,S., Altafi,H., Araujo,R., Conn,L., Conway,A.B.,
Gonzalez,A., Hansen,N.F., Huizar,L., Kremenetskaia,I., Lenz,C.,
Li,J., Liu,S., Luros,S., Rowley,D., Schwartz,J., Toriumi,M.,
Vyotskaia,V., Yu,G., Davis,R.W., Federspiel,N.A., Theologis,A. and
Ecker,J.R.
Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
I
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 132699)
Ecker,J.R.
AUTHORS Direct Submission
TITLE Submitted (25-FEB-1999) Arabidopsis thaliana Genome Center,
JOURNAL Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
```

REFERENCE AUTHORS TITLE JOURNAL	3 (bases 1 to 132699) Ecker,J.R. Direct Submission Submitted (07-MAY-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	CDS	GIAGKSKAEFOAPCRRTKYPREIPLPWYRSALPQAMAGFLPFPFSAIYIELYVIFAS VWGHRIYIYSILFIVFIIITAFITVALTYFQAAEDHOMWRSFLCGSGTGLFI YAYCLYYIYARSOMSGFMQTSFFFGYMACICYGFFLMLGTVGFRALLFVRHYRSIK CE"
REFERENCE AUTHORS	4 (bases 1 to 132699) Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On May 7, 1999 this sequence version replaced gi:4731042. Location/Qualifiers 1. 132699 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" complement(join(144..278,360..491,576..673,814..1435, 1749..1867,2161..2310,2441..2565,2698..2742,2837..2902, 2979..3031,3117..3234,3314..3415,3498..3655,3741..3818, 3894..3985,4054..4104,4138..4420)) /note="similar to Ne/H antiporter dbj BAA89487.1" /codon_start=1 /evidence=not_experimental /product="F10B6.1" /protein_id="AAF79251.1" /db_xref="GI:8778242"	CDS	8842..9714 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.4" /protein_id="AAF79218.1" /db_xref="GI:8778209" /translation="MDLSREVDFOFIKETIDHSGLPLSDVQLKLTAEESQRRLRE QYLSVSLKEKEKVIDLVRSEASMAQSLKKFVEENOKLGSEREDLVNCKWEKEC FLYHQDESLSMGNETDARAEARSRELEEVKMSDEIKRSIESEEDLVDSIL ASFVQDESISLGRIFLEANSEDKCCTELLRKWDQKPTQKFVSLVSMVKRIEKK ECLIMLAKAEQVELVSEQNRELDRENKFLRQCSAERSHSGSNKFNKRSIKMSSP IEKRIELSSQEEFD" 11067..11906 /note="unknown protein; similar to EST gb AI995472.1" /codon_start=1 /evidence=not_experimental /product="F10B6.5" /protein_id="AAF79219.1" /db_xref="GI:8778210" /translation="MDDDGFRNMGYEPAAATFGNGLQMLSTIDNTKPFPLGRDP NLMGPNQSYHHQEPPIHMSYNWQKDKFNNMLPYTTATPNYGLPETSSAPSHQ MNLHHLOTEENPVKLEETIVVOTKKRKTNAKAGSTPAKKPRKPKDENNNNNNT NVTRVKPAKSVDLVINGVSMDSGLPVICTCTGAPQCYRGCGQWQACCTTNTS MHPLPNSTKRGARISGRKMSQGAFFKVKLEKLASDGFNGNPIDLKSHWARHGNTKFV TIR"
TITLE JOURNAL	complement(13762..14781) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.6" /protein_id="AAF79250.1" /db_xref="GI:8778241" /translation="MKLWRAAVVLKDGPSLIAADDILTAAVVKATSHDELSIDTESA QFYIRHVLSSPSSLKPLVSLTSSRVKTRSWAVALKMLMHGFFLCKSTVAESIGRL PFDLSQEGENSRIMSGGNFLPVRAVFAFLDRRSILFHDGNHRYNEESSVIRLV IIRKMOJLDSLIRIKPIGEMMIPVINEAMENVSEIMELYGICRIEAEVLPNVIS KIKTEADLALKIVAKSMKGGLKKYFEFCDKLGVNAQEIPIPNFVRIPDADVIHDE LVRTAMSSSESAEETAEIEEEEEIEETKLSLDLITLDHNEEPAPSPRVVVVDIP DLFSF"	CDS	16027..16533 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.7" /protein_id="AAF79220.1" /db_xref="GI:8778211" /translation="MSTCYVRRCMRNHAALGSIADGCREYSQSTGDLVACGCH RSHRRIDVSSQINHTREFPLRLRQLARKWTAEERNEEEDDEETESTEK MTVORRRKSKFTAEQREAMKDYAAKGLWTLKDKALREERFVCEGIVTRYHPTWV NNKKYH"
COMMENT FEATURES source	join(17635..17815,17901..18019,18141..18194, 18300..18406,18547..18611,18742..18950)) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.8" /protein_id="AAF79249.1" /db_xref="GI:8778240" /translation="MEFKAYFILYWFQFPFLQCEPVIVFLCFQMICKKKQRIFLFW YPEELHQPLTPMVIWVLFIRIWRWFDQVCKSSTGTGDSLVSGAEGALGFNE IVRPYEVKSLGSEKSLCPKSGKSVIFGQALKVGTSLTFVGEAVRDKAGNLMTQSK EQSLIVFESEESFDEMVMNMKSELCAVILAKIFGSIATAVAVVYGVYDVARVLLPFFV WKKDLGSSSSQSSSDSDTDSDF" complement(join(20071..20409,20499..20573,20648..20721, 20800..20887,20957..21108,21188..21356,21448..21906, 22006..22275,22367..22570,22840..22917)) /note="hypothetical protein" /codon_start=1	CDS	18027..18533 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.9" /protein_id="AAF79251.1" /db_xref="GI:8778242" /translation="MSTCYVRRCMRNHAALGSIADGCREYSQSTGDLVACGCH RSHRRIDVSSQINHTREFPLRLRQLARKWTAEERNEEEDDEETESTEK MTVORRRKSKFTAEQREAMKDYAAKGLWTLKDKALREERFVCEGIVTRYHPTWV NNKKYH"
CDS	complement(join(17635..17815,17901..18019,18141..18194, 18300..18406,18547..18611,18742..18950)) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.10" /protein_id="AAF79250.1" /db_xref="GI:8778240" /translation="MEFKAYFILYWFQFPFLQCEPVIVFLCFQMICKKKQRIFLFW YPEELHQPLTPMVIWVLFIRIWRWFDQVCKSSTGTGDSLVSGAEGALGFNE IVRPYEVKSLGSEKSLCPKSGKSVIFGQALKVGTSLTFVGEAVRDKAGNLMTQSK EQSLIVFESEESFDEMVMNMKSELCAVILAKIFGSIATAVAVVYGVYDVARVLLPFFV WKKDLGSSSSQSSSDSDTDSDF"	CDS	18027..18533 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.11" /protein_id="AAF79251.1" /db_xref="GI:8778242" /translation="MSTCYVRRCMRNHAALGSIADGCREYSQSTGDLVACGCH RSHRRIDVSSQINHTREFPLRLRQLARKWTAEERNEEEDDEETESTEK MTVORRRKSKFTAEQREAMKDYAAKGLWTLKDKALREERFVCEGIVTRYHPTWV NNKKYH"
CDS	complement(join(17635..17815,17901..18019,18141..18194, 18300..18406,18547..18611,18742..18950)) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.12" /protein_id="AAF79250.1" /db_xref="GI:8778240" /translation="MEFKAYFILYWFQFPFLQCEPVIVFLCFQMICKKKQRIFLFW YPEELHQPLTPMVIWVLFIRIWRWFDQVCKSSTGTGDSLVSGAEGALGFNE IVRPYEVKSLGSEKSLCPKSGKSVIFGQALKVGTSLTFVGEAVRDKAGNLMTQSK EQSLIVFESEESFDEMVMNMKSELCAVILAKIFGSIATAVAVVYGVYDVARVLLPFFV WKKDLGSSSSQSSSDSDTDSDF"	CDS	18027..18533 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F10B6.13" /protein_id="AAF79250.1" /db_xref="GI:8778240" /translation="MEFKAYFILYWFQFPFLQCEPVIVFLCFQMICKKKQRIFLFW YPEELHQPLTPMVIWVLFIRIWRWFDQVCKSSTGTGDSLVSGAEGALGFNE IVRPYEVKSLGSEKSLCPKSGKSVIFGQALKVGTSLTFVGEAVRDKAGNLMTQSK EQSLIVFESEESFDEMVMNMKSELCAVILAKIFGSIATAVAVVYGVYDVARVLLPFFV WKKDLGSSSSQSSSDSDTDSDF"







```

/clone="BwA6-II"
/note="synonym: Vouacapoua americana"
repeat_region 1..524
/note="microsatellite"
/rpt_type=tandem
/rpt_unit=ga
BASE COUNT 144 a 90 c 182 g 108 t
ORIGIN

```

```

alignment_scores:
  Quality: 51.00      Length: 16
  Ratio: 3.643       Gaps: 0
  Percent Similarity: 87.500  Percent Identity: 62.500
alignment_block:
US-09-485-529-104 x AF224770/rev ..

```

Align seg 1/1 to reverse of: AF224770 from: 1 to: 524

```

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSeraspMet 16
|||||
134 GACGCTCTCTCGCGCGCGGATACAGGTGAGTCTCGGAGCTC 87

```

seq\_name: gb\_htg:AC015886

seq\_documentation\_block:

LOCUS AC015886 188716 bp DNA HTG 07-MAY-2001  
DEFINITION Mus musculus clone CT7-378P20, WORKING DRAFT SEQUENCE, 26 unordered pieces.

ACCESSION AC015886

VERSION AC015886.4 GI:13959257

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 188716)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Mus musculus chromosome, clone CT7-378P20

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 188716)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B., Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,K., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,K., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wymann,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

TITLE Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome

JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On May 7, 2001 this sequence version replaced gi:7230250.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L737

Center clone name: 378\_P20

----- Summary Statistics

Sequencing vector: M13; M77815; 82% of reads

# FEATURES

source

Sequencing vector: Plasmid; n/a; 18% of reads  
Chemistry: Dye-primer-amerham; 84% of reads  
Chemistry: Dye-terminator Big Dye; 16% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 175387 bases at least Q40  
Consensus quality: 181747 bases at least Q30  
Consensus quality: 184307 bases at least Q20  
Insert size: 173000; agarose-fp  
Insert size: 186216; sum-of-contigs  
Quality coverage: 4.8 in Q20 bases.  
NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

Contig	Start	End	Length
1	1020	contig of 1020 bp	in length
1021	1120	gap of 100 bp	
1121	2393	contig of 1273 bp	in length
2394	2493	gap of 100 bp	
2494	3639	contig of 1146 bp	in length
3640	3739	gap of 100 bp	
3740	4886	contig of 1147 bp	in length
4887	4986	gap of 100 bp	
4987	6130	contig of 1144 bp	in length
6131	6230	gap of 100 bp	
6231	7448	contig of 1218 bp	in length
7449	7548	gap of 100 bp	
7549	8847	contig of 1299 bp	in length
8848	8947	gap of 100 bp	
8948	9970	contig of 1023 bp	in length
9971	10070	gap of 100 bp	
10071	11323	contig of 1253 bp	in length
11324	11423	gap of 100 bp	
11424	12572	contig of 1149 bp	in length
12573	12672	gap of 100 bp	
12673	14167	contig of 1495 bp	in length
14168	14267	gap of 100 bp	
14268	15746	contig of 1479 bp	in length
15747	15846	gap of 100 bp	
15847	17416	contig of 1570 bp	in length
17417	17516	gap of 100 bp	
17517	18799	contig of 1283 bp	in length
18800	18899	gap of 100 bp	
18900	20317	contig of 1418 bp	in length
20318	20417	gap of 100 bp	
20418	22467	contig of 2050 bp	in length
22468	22567	gap of 100 bp	
22568	25820	contig of 3253 bp	in length
25821	25920	gap of 100 bp	
25921	29289	contig of 3369 bp	in length
29290	29389	gap of 100 bp	
29390	34348	contig of 4959 bp	in length
34349	34448	gap of 100 bp	
34449	44539	contig of 10091 bp	in length
44540	44639	gap of 100 bp	
44640	55425	contig of 10786 bp	in length
55426	55525	gap of 100 bp	
55526	72079	contig of 16554 bp	in length
72080	72179	gap of 100 bp	
72180	85873	contig of 13694 bp	in length
85874	85973	gap of 100 bp	
85974	118623	contig of 32650 bp	in length
118624	118723	gap of 100 bp	
118724	151284	contig of 32561 bp	in length
151285	151384	gap of 100 bp	
151385	188716	contig of 37332 bp	in length.

Location/Qualifiers

1..188716

/organism="Mus musculus"

/db\_xref="taxon:10090"

```
/clone="CT7-378P20"
/clone_lib="CitbcJ7 Mouse BAC"
1. 1020
/note="assembly_fragment"
1121..2393
/note="assembly_fragment"
2494..3639
/note="assembly_fragment"
3740..4886
/note="assembly_fragment"
4987..6130
/note="assembly_fragment"
6231..7448
/note="assembly_fragment"
7549..8847
/note="assembly_fragment"
8948..9970
/note="assembly_fragment"
10071..11323
/note="assembly_fragment"
11424..12572
/note="assembly_fragment"
12673..14167
/note="assembly_fragment"
14268..15746
/note="assembly_fragment"
15847..17416
/note="assembly_fragment"
17517..18799
/note="assembly_fragment"
18900..20317
/note="assembly_fragment"
20418..22467
/note="assembly_fragment"
22568..23820
/note="assembly_fragment"
25921..29289
/note="assembly_fragment"
29390..34348
/note="assembly_fragment"
34449..44539
/note="assembly_fragment"
44640..55425
/note="assembly_fragment"
55526..72079
/note="assembly_fragment"
72180..85873
/note="assembly_fragment"
85974..118623
/note="assembly_fragment"
118724..151284
/note="assembly_fragment"
151385..188716
/note="assembly_fragment"
52371 a 41183 c 41183 g 51477 t 2502 others
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 11
  Ratio: 4.455        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 90.909

alignment_block:
US-09-485-529-104 x AC015886/rev ..
Align seg 1/1 to reverse of: AC015886 from: 1 to: 188716
2 GtUeuLeuAlaAlaLeuGlyTyrLysValArg 12
|||||
121315 GAGCTTTGGCAGCCTAGGTATAGGTGAGA 121283

seq_name: gb_pr:AL355796
```

```
seq_documentation_block:
LOCUS      AL355796      152086 bp      DNA      PRI      23-DEC-2000
DEFINITION Human DNA sequence from clone RP11-46B11 on chromosome 6, complete
sequence.
ACCESSION  AL355796
VERSION    AL355796.11 GI:11991379
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 152086)
AUTHORS    Almeida, J.
TITLE      Direct Submission
JOURNAL    Submitted (23-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT    On Dec 24, 2000 this sequence version replaced gi:11990033.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-46B11 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-46B11. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-46B11 is at 1 in this sequence. The
true left end of clone RP3-508D13 is at 151987 in this sequence.
The true right end of clone RP11-454P17 is at 95829 in this
sequence.

FEATURES
             Location/Qualifiers
             1..152086
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="6"
                /clone="RP11-46B11"
                /clone_lib="RPCI-11.1"
             1..1193
                /note="L1 repeat: matches 2364..3549 of consensus"
             1194..1456
                /note="L1PA7 repeat: matches 5881..6143 of consensus"
             1457..2384
                /note="L1 repeat: matches 1432..2364 of consensus"
             2182..3056
                /note="L1PAa repeat: matches 235..1639 of consensus"
             3081..4141
                /note="L1PAa repeat: matches 1537..412 of consensus"
             4146..7688
                /note="L1PA13 repeat: matches 1832..5346 of consensus"
             7689..7939
                /note="L1PA9 repeat: matches 5909..6163 of consensus"
             7940..8737
                /note="L1PA13 repeat: matches 5346..6156 of consensus"
```



seq_documentation_block:					
LOCUS	AC007504	125021 bp	DNA	PLN	11-JUN-2001
DEFINITION	Arabidopsis thaliana chromosome I BAC F13F21 genomic sequence, complete sequence.				
ACCESSION	AC007504				
VERSION	AC007504.3	GI:5430745			
KEYWORDS	HTG.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 125021)				
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 125021)				
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-MAY-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
REFERENCE	3 (bases 1 to 125021)				
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-JUN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
REFERENCE	4 (bases 1 to 125021)				
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P., Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S., Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V., Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-JUL-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
COMMENT	On Jul 9, 1999 this sequence version replaced gi:5019265. e-mail for correspondence: arabsequence.stanford.edu Genes with similarity to proteins in the databases are described as 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <a href="http://compbio.ornl.gov/section/index.html">http://compbio.ornl.gov/section/index.html</a> ), GENSCAN (Chris Burge, & A.Salamov, Sanger Centre, <a href="http://genomic.sanger.ac.uk/">http://genomic.sanger.ac.uk/</a> ), NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <a href="http://www.cbs.dtu.dk/NetPlantGene.html">http://www.cbs.dtu.dk/NetPlantGene.html</a> ) and eMotif (Nevill-Manning, C.G. Wu, T.D. & Brutlag, D.L., <a href="http://motif.stanford.edu/projects.html">http://motif.stanford.edu/projects.html</a> ). Location/Qualifiers 1. .125021 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="I" /map="ni208" /clone="F13F21" complement(4079. .5020)				
FEATURES	source				
	gene				

```

/gene="F13F21.1"
complement(join(4079. .4465,4631. .5020))
/gene="F13F21.1"
/feature="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43146.1"
/db_xref="GI:5430746"
/translation="MTNLGSSYPTANLYFMQVWKIECWLRHAFESVDETCQMVET
MTLKFVWEYSIDILAAVDFPRLKLCLEYCFSTLDRLTSKSLAHVRSKLYKLF
KAYKRPSSTSSQVETLEEDIPAGYGEVAFYVQKVGSSGKSELDIYLGEPFLDMA
AFRIFNLAYWKNDSQVETLEEDIPAGYGEVAFYVQKVGSSGKSELDIYLGEPFLDMA
ENIQALICTNRNLGFPKGEVEVEKEEKEEKEE"
6527. .10142
/gene="F13F21.2"
join(6527. .6667,6772. .6859,7159. .7311,7597. .7757,
7840. .8514,8612. .8939,9020. .9240,9369. .9546,9658. .10142)
/gene="F13F21.2"
/feature="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43147.1"
/db_xref="GI:5430747"
/translation="MSENTKVEAKRVFVIGAGCNVNVNWSGASLYSFGAQNAAVAF
CPKTAQILTLPGHKASVNCVTHWLPSTKFAKAKKLDQRYLSSDGSIGIILWELSTL
NDVKFYSRSLDSFFAVLSLIPKECKVCLDSICVDTKAIYVLSLAEUQPMGRF
ALALGLDNKIKLYSGERTKFTSVCKELGHTDWIRSLDFSLPHRTTEIPNSLTHLS
SSQDKVIRIKVLVGVDSWRREITLASVIEGPFVSGTFTQYQISVSVLIGHEDWY
YSVEMQPVDFIDGRLVHQPLSILASMDKTMWIRPEKTKTGWNVNVVCGVGLSHC
ALGPYGHWSNLSILAHGYGGAFHLWRNVSSKSESNQKQVPSGHFAAVDVTW
ARTGEYLLSVSDQTTVFSAKNDEGNEADHHLARPQVHGHINDVAVVQVGGK
NHRVSGAEKRVVFAPLSLNLNCTAGGEGSPFEDIQADVQLGANSALGLS
QKPYLHSSSEPERNGGGLDFTFVPEAPAEKPEPDEDQAFHETLMPESHKLY
GHNEFLSLCSHDHGNLVASCKAQASMAEILWVEGTWKAVGRQASHSLTVHLEF
SYDDTLSSVSRDRHVSFVSTQRTDNGEVSHKMAKVEAHKRITWAGSNVFPGHQFAT
SSRDTKVINSVENDARIKQILLVPPGSSVTAVAWTGLDRNEKSGCVAVGMESSLIE
LSNVKIIETEGTATTAALALRLEPFMCHVSANRLAWRPEKCESQSLRWLISCGD
DNCVRVFNFKP"
complement(10370. .11380)
/gene="F13F21.3"
complement(10370. .11380)
/gene="F13F21.3"
/feature="97% identical to geranyl geranyl pyrophosphate
synthase [Arabidopsis thaliana] gi|2578822."
/codon_start=1
/product="geranyl geranyl pyrophosphate synthase"
/protein_id="AAD43148.1"
/db_xref="GI:5430748"
/translation="MRPRYSLLSAMRLIRPSNRRLSIASSDSEFTSYMKNAKASIN
KALNSIPLCNFVPLWEPVLEVKAMRYLLPGCKVRPMLCLVACELVGGQSTAM
PAACAVEMHAASLIDDLPCMDSDSLRRGPTNKKVFGERTSILASNALSLAVKQT
LASTSLGVTSEVRLVAVOEMARAVGTGLVAGAADLAGERMSFKNEDDEURYLELMH
VHKTAVLVEAAVYGAINGGSDSEIERLKSARCVCGLMFQVMDVLDLDTKSSSELHG
TAGKDLITGKLTYPKMGVDNAREYAKRLNREAQHLQGFSDKVPVLLSLADYIVKR
QN"
complement(12218. .14173)
/gene="F13F21.4"
complement(join(12218. .12318,12359. .12418,12518. .12695,
12784. .12898,12987. .13143,13236. .13410,13500. .13602,
13902. .14173))
/gene="F13F21.4"
/feature="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43149.1"
/db_xref="GI:5430749"
/translation="MWSDSLVTOLRETLRSSDLETTTPASVRQLEVYFGVELTDKK
AFVREQIDAFLESALLLESKEPEQEECDQNDDEEGSDDDKTELPAVKAKKGGGF
NKICQLSPLEKFLGTSQARTVYVKKWYIRHDLQDPTNRRNLCDLSLHSLFRV
KTINFMQNKALAKHIWALNDGCFKNKEEDVDETSGERDEKQKIEALEENNEE
SREEDRSVRRKKRKKPAKSEKPKKGGFTKVCSLSPQLAQFTPOLARTVEV
KMLWKYIKENNLQPSDKRTICDESLSLFPVESINNFQNKOLAKHINPLVQDEEA
GTTNDPEKGQKMKMETDEVKNIKSSTFLNVIDNESNEEKATSSRIKTEE"
complement(15762. .16883)
/gene="F13F21.5"
complement(join(15762. .15977,16061. .16492,16809. .16883))

```

```
/gene="F13F21.5"
/notes="Unknown Protein; Location of ests VBQD12
(gi|757594 and gi|757593)"
/codon_start=1
/protein_id="A043150.1"
/db_xref="GI:5430750"
/translation="MAMSHLFLSSPQSSALRLHSTTOFTLSYKNNKDCFSQANE
AKVRSRLCRAIHMSGSHGQPKKLNFDNLLRRTKHVNDSPQVFKFPFNRAFGN
FIQLVDLAISSVVKLFVPLVLAIVSISEMSYCAHERKALVFPFVIGVAVGLQET
ALKISPRLKEAVPQHLTAMMEFTLIKLPGPYPYWGRLLYPHFANGVLLRALMSMF
FHYKTRNTSCNPLQHSLETE"
21730..21921
/gene="F13F21.6"
21730..21921
/gene="F13F21.6"
/notes="Unknown protein; Location of ests 164G18T7
(gi|2764114, 114G10T7 (gi|2597578), and 114G11T7
(gi|2597579)"
/codon_start=1
/protein_id="A043151.1"
/db_xref="GI:5430751"
/translation="MVTLMKEICIELLKLTVDFVAVAESIEVAFRHRPPVQYSSV
MGRSNTAVPIPLVGL"
23803..26346
/gene="F13F21.7"
23803..26346
/gene="F13F21.7"
/notes="Hypothetical Protein"
/codon_start=1
/protein_id="A043152.1"
/db_xref="GI:5430752"
/translation="WERPPGCFILLISYTVVATFDDPSPENADLTKELEKCF5
INKVPLKFNDRKRAVIALQAMKKAISDPFKTANNVGSVCYNGVYCAPALD
DLSLVAVGVQDLNADIAHLPPELGLMTDLALFHINSNRFCGIIPKLSKLAMTEF
DVSNNRVGDFEVLSSWPLFLDLRYNEFEFGSPSEIFDKDLDAIFLNNRFSVI
PTIGKASVGVTFANNKSGCIPKSIIGNKNLNEIVFTGNLTGCFPEIGLLNVT
VFDAGKNGVLPSTPLSGLASVQDLDSHNKLTGFVKDLKPLNLD5FKFSYNFVN
GAESCVPGRNNGKDFDFTNGLONRPSOKPAKQCLPVVSRPVDCSKDKCSGNSGS
SPSPNPTSEPKSPKPEVMPKPSDSSKPEPKPEQSPKPEPKHSPKPEPEN
KIELPKQKSPKQSPKSPKSPKPEQSPKPEQSPKPEQSPKPEQSPKPEQSPK
DPYDASVPKRRSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PPPVVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PPPVVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PVOQPTPVQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SSESQSPQATPPILEPVHATPNSKPVQSPSPSPSPSPSPSPSPSPSPSPSP
SVPSSTSDTSIPFPENDDDDGDFVLPFHIGQYASPPPPPEQY"
28449..29313
/gene="F13F21.8"
Join(28449..28680,28766..28965,29065..29313)
/gene="F13F21.8"
/notes="Unknown Protein"
/codon_start=1
/protein_id="A043153.1"
/db_xref="GI:5430753"
/translation="MQMDSAQONKRRARLFEDPELKDVIYPSNPSTEPVKNKGY

alignment_scores:
  Quality: 47.00      Length: 13
  Ratio: 3.615       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 69.231

alignment_block:
US-09-485-529-104 x AC007504/rev ..
Align seq 1/1 to reverse of: AC007504 from: 1 to: 125021
2 GluLeuLeuAlaAlaLeuGlyTyrTlysValArgAlaSer 14
|||||
112481 GAGCTACTAGCATCCTTAGGTATATAACAAACTTCC 112443

seq_name: gb_htg:AC034100
seq_documentation_block:
```

```
LOCUS AC034100 179481 bp DNA HTG 13-JUL-2001
DEFINITION Mus musculus chromosome 14 clone RP23-46517 strain C57BL6/J,
WORKING DRAFT SEQUENCE, 14 unordered pieces.
ACCESSION AC034100
VERSION AC034100.8 GI:14717146
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 179481)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
JOURNAL Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
REFERENCE Gordon,M., Goltz,J.S. and Kucherlapati,R.
AUTHORS High Throughput Mouse Sequencing
TITLE Unpublished
JOURNAL 2 (bases 1 to 179481)
AUTHORS Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
JOURNAL Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
REFERENCE Gordon,M., Goltz,J.S. and Kucherlapati,R.
AUTHORS Direct Submission
TITLE Submitted (04-APR-2000) Department of Molecular Genetics, Albert
JOURNAL Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
COMMENT Bronx, NY 10461, USA
On Jul 13, 2001 this sequence version replaced gi:10937953.
-----Genome Center
Center: Albert Einstein College of Medicine
Center Code: AECOM
Web site:
http://sequence.aecom.yu.edu/cgi-
bin/ws.exe/mouseDB/mouseSEQ/mouseseqtable.hts
Contact: htgs@sequence.aecom.yu.edu
-----Summary Statistics
Center project name: ABL
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 175183 at least Q20
*Consensus quality: 174588 at least Q30
*Consensus quality: 173495 at least Q40
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 179221 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 9.4 x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 50762: contig of 50762 bp in length
* 50763 50782: gap of unknown length
* 50783 92572: contig of 41790 bp in length
* 92573 92592: gap of unknown length
* 92593 122766: contig of 30174 bp in length
* 122767 122786: gap of unknown length
* 122787 151381: contig of 28595 bp in length
* 151382 151401: gap of unknown length
* 151402 168067: contig of 16666 bp in length
* 168068 168087: gap of unknown length
* 168088 174488: contig of 6401 bp in length
* 174489 174509: gap of unknown length
* 174509 175363: contig of 855 bp in length
* 175364 175384: gap of unknown length
* 175384 175685: contig of 282 bp in length
* 175686 175685: gap of unknown length
* 175686 176803: contig of 1118 bp in length
* 176804 176823: gap of unknown length
* 176824 177025: contig of 202 bp in length
* 177026 177045: gap of unknown length
* 177046 177244: contig of 199 bp in length
```



```

gene
3087. .4040
/locus="m1r1247"
CDS
3087. .4040
/locus="m1r1247"
/codon_start=1
/transl_table=11
/product="esterase"
/protein_id="BAB48665.1"
/db_xref="GI:14022055"
/translation="MFSLKSHVSVFLVRSKKAFFSPENLRWIAAAKRTDHHPPA
SGHQIHTVTPDVGSPVYIAPKAGEQRILYLHGGAYFEITPYWHLLADMDRL
LGHQITVPIAPHDFHAFMGVQVYRQMLDETAEDIVFMGDSAGGNMAYVLRMM
AAEGLPRAHVLSPLGDMLSNPKLFEARNRDPWLGAGLEAIRMYSGIDRSD
WHISPTYGVSVLPTLLLTGSHDLLTPDNLIPAOKARAGVEVEVHEEGMHPANPL
IDMPEARROHIVAFLGNRSPVRQDGKTRFOAPSAAAE"
complement(4062. .4961)
/gene="m1l1248"
CDS
complement(4062. .4961)
/gene="m1l1248"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="BAB48666.1"
/db_xref="GI:14022056"
/translation="MNAIPLNHPDLLDLVLRTEFVAIAETGSGFTTAANAVERTPSAVS
MOIKLEDILGRSVFARDARSVSLTDTGEMLLGYARLLISINREVSKFLLPIGVV
RLGSPDYGERVLPVHLKRPQAQSPHIAVDVTIDQSNLRMRDDRALDITLLTNSYK
TSALEAVLLETPFVWAGAGGCAHLREPLPVSLWEEGCARAGALEALGREGNRYR
AYMSATAGORAIADLAVAPLPRSPFLGNDMVYELCPKGMPPDIGTYNLAMVVPADAS
APVKAVADHIRATFEVRETGF"
5113. .5379
/gene="m1r1249"
CDS
5113. .5379
/gene="m1r1249"
/codon_start=1
/notes="unknown protein"
/transl_table=11
/protein_id="BAB48667.1"
/db_xref="GI:14022057"
/translation="MGLKHAACKTSQIKRNVPGPAPKENEMSLSLGRIAFEFSAA
RARYQTEAIRSLPIELQKDIGWPEASDTKTGTNRNGVSGWAGAK"
6136. .7137
/gene="m1r1251"
CDS
6136. .7137
/gene="m1r1251"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="BAB48668.1"
/db_xref="GI:14022058"
/translation="MNAIKHPKRSFVFLVPDFTMIATATALDPLRSANRMLGYEAY
RWKLASDGKPVRSAGVECAVNTSLEERKNWAGDPNPMALVCSGVNVERVQNSA
FALREYNRGVAGGLCTGAHTLAAAGLISNKRCAIHWENLFGSEAFKAVFADL
FEVDQNIYCAGTAALDMMKLLIGDDFDESILNRVCEQVLTDRVSRPTDRQLPLRA
RLGVQSKVITIELMEGNLSEPLSLEIETHVDLSKROELRFRTEMGRSPARYYLE
IRLDRARHLLIQSSMPVVEVAVACGFSVASHFSKCYRELYARSPQQRVDRKQLLAA"
complement(7153. .7956)
/gene="m1l1253"
CDS
complement(7153. .7956)
/gene="m1l1253"
/codon_start=1
/transl_table=11
/product="ATP-binding protein of ABC transporter"
/protein_id="BAB48669.1"
/db_xref="GI:14022059"
/translation="MEPGVSKLLIEGVSRFAGVGLPVKALMPIDLVAANDFTTI
LQPSGGCKSTLLRIVAGLEAPSRGVLLDDKAVTRPGDPDMGVFQSYTLPPLVTSEN
IAGFLRGNWAPKERRDDIVASVYDLVLKGFENHPKQLSGCGQQTATARALNDPE
ILLDPEFGALDQRLGMLQELLGIWERKKVLFVTHDIEEAFIMASVIVMTARP
GSISKDVAIDLPHRYITKTSPEFSALKARLUTDIRVEAMRTAQAPQ"
complement(7941. .8720)
/gene="m1l1254"

```

```

CDS
complement(7941. .8720)
/gene="m1l1254"
/codon_start=1
/transl_table=11
/product="permease protein of ABC transporter"
/protein_id="BAB48670.1"
/db_xref="GI:14022060"
/translation="MKMRPLHPVSPGLRTGLGISFFFLFIAFWAIFLGGHVNRIFLA
DPLSLRDGWRLLVEDRFWLDILITWRVGGFVLASVAVPTICIAMGANKPVEAFLE
PFVSPARYLPASAFIPLLILWAGIGEKLILYFVGSVFOIILIVAKVGSRRDLVE
AATLTSTNGNIVKRVIMPANAEIATETLLVLGNWAVTIVIAELIGSSSGIGITMIIN
SQSLRATGQIFIGIIVIGLIGLSDFAFKAFNRWLPWPSLA"
complement(8802. .9752)
/gene="m1l1256"
CDS
complement(8802. .9752)
/gene="m1l1256"
/notes="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48671.1"
/db_xref="GI:14022061"
/translation="MNRIVAFATLTFAGLSAPVMAATSVTIGISGWTGFAPLTLAK
QAGLEKHGLDVTLLKVPQASRPLATASGDLCAATVETWLVNMSGVTTKQIFQLD
KSYGADGIVVRNDIKTVADLKGNVASSAPGTSPTFLAWLNKNGMSTKDVTVNLE
PDAQAQFAGQNDAAVTYEPFISAVRDKADQGHILATTLTDYPMVLDTVGTCTPEFLKA
NPDAAKALADSYFEALDLIKKDPKSYEIMGADVKSAREFEDSAKYLKWADKADNQ
FFTKEDQESKTAGALLQMLIKEAPDVATLADTSAVAN"
complement(9963. .11000)
/gene="m1l1257"
CDS
complement(9963. .11000)
/gene="m1l1257"
/notes="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48672.1"
/db_xref="GI:14022062"
/translation="MSISARIPRSKPSVATSRFTRTSSRAADRSGGPRDLLQKP
GFEPGLFKQVVFAYPRGMSALSRARHLLSGCAALMLWLPMSRADPVDVVELV
ADVSLSMSADELEIQHGYAALTDHNVLOATADGAHKIATVYVSWAGTTQWRVIV
PWTVIANRDAERVVAQLSAQPPNSARTSISGALEFGSLFAESGEGTKRVIDISG
DGNPNQGPVNLTRDGVROGIVINGLPLMRGLSGAYDNDLDRIYSDCVIGGPGA
FMTPVNDQTFQPEAIRKRLVLELAPSPQWAAAEADHPFVLTQDKPAADCCQIGEM
WRNRNWLDSR"
complement(10906. .13467)
/gene="m1l1258"
CDS
complement(10906. .13467)
/gene="m1l1258"
/codon_start=1
alignment_scores
Quality: 47.00 Length: 16
Ratio: 3.357 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 62.500
alignment_block
US-09-485-529-104 x AP002997/rev ..
Align seg 1/1 to reverse of: AP002997 from: 1 to: 329709
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetala 17
|||||
249258 GAGGTGCTGCACAACTCGGTTACAAGGTGCAAGGTCGACAGGCC 249211
seq_name: gb_pat:AX005855
seq_documentation_block:
LOCUS AX005855 309 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 64 from Patent WO9909174.
ACCESSION AX005855
VERSION AX005855.1 GI:9928850
KEYWORDS bread wheat.
SOURCE
ORGANISM Triticum aestivum

```



Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 309)  
AUTHORS Harberd,N.P. and Peng,J.  
TITLE Genetic control of plant growth and development  
JOURNAL Patent: WO 9909174-A 64 25-FEB-1999;  
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)

FEATURES Location/Qualifiers

1..309

/organism="Triticum aestivum"

/db\_xref="taxon:4565"

9

/note="n is any nucleotide"

46

/note="n is any nucleotide"

50

/note="n is any nucleotide"

72

/note="n is any nucleotide"

95

/note="n is any nucleotide"

131

/note="n is any nucleotide"

141

/note="n is any nucleotide"

154

/note="n is any nucleotide"

160

/note="n is any nucleotide"

172

/note="n is any nucleotide"

176

/note="n is any nucleotide"

214

/note="n is any nucleotide"

222

/note="n is any nucleotide"

47 a 102 c 102 g 45 t

BASE COUNT 47 a 102 c 102 g 45 t

ORIGIN

alignment\_scores:  
Quality: 46.00 Length: 9  
Ratio: 5.111 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-485-529-104 x AX005855

Align seg 1/1 to: AX005855 from: 1 to: 309

8 GtYtYrlysvaIargAlaSerAspMet 16

|||||

18 GGTACAAAGTGGCGCCCTCCGACATG 44

seq\_name: gb\_ba:RPXX02

seq\_documentation\_block:

LOCUS RPXX02 312430 bp DNA BCT 12-NOV-1998

DEFINITION Rickettsia prowazekii strain Madrid E, complete genome; segment

2/4.

ACCESSION AJ235271 AJ235269

VERSION AJ235271.1 GI:3868717

KEYWORDS complete genome.

SOURCE Rickettsia prowazekii.

ORGANISM Rickettsia prowazekii.

Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

Rickettsiaceae; Rickettsieae; Rickettsia; typhus group.

1 (bases 1 to 312430)

REFERENCE 1 (bases 1 to 312430)

Anderesson,S.G., Zomorodipour,A., Andersson,J.O.,

Sicheritz-Ponten,T., Alsmark,O.C., Podowski,R.M., Naslund,A.K.,

Eriksson,A.S., Winkler,H.H. and Kurland,C.G.

The genome sequence of Rickettsia prowazekii and the origin of

mitochondria

Nature 396 (6707), 133-140 (1998)

99039499

2 (bases 1 to 312430)

Andersson,S.G.E.

Direct Submission

Submitted (11-NOV-1998) S.G.E. Andersson,

Siv.Andersson@molbio.uu.se, Dept. of Molecular Biology, University

of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN

On Nov 13, 1998 this sequence version replaced gi:3860788.

Location/Qualifiers

1..312430

/organism="Rickettsia prowazekii"

/strain="Madrid E"

/db\_xref="taxon:782"

76..1830

/gene="RP226"

76..1830

/gene="RP226"

/codon\_start=1

/transl\_table=1

/product="unknown"

/protein\_id="CAA14689.1"

/db\_xref="GI:3860789"

/translation="MLNKLCDILFLINLLVTVQGYASPPPLPSPIIAVDTDKNI

STNSISFFEFKQFFSKQKKNISQHEQOTKAIHQESQIDISRELNEQSEPEFI

DFGSTILPSVASYIDSKAEYENSTNLAASTODIQVKQEPDPSASEPIDGNTK

FTSATNHEMYKEAVSSNDKTNLSNIIITNPVSPVISIPTAQDVNVVPSQSVQIY

KPTNLTSIRNPIDNHHHTDLNKEKLESTISNMTTITNNVSPVSDIQTITLNT

VPTAEHVSVQTSVTHSNQHSQAPITPISINTPVTSSVTLRATESVPIINNSOEIF

VSESESTKQDWTYPIMPVLVVDPNKQSKPLALEKNNNDIIINQAESHVSSEN

TIQKNDKVNNAATSESTKEFYKNETQMLFPDDDIVLGKLTEDATLQMDMHGVIKLF

QKKEWTANAERKLVESTKYDDDKNDINKNDIYANLSYSAVDNAFRAVENKNLFEIR

ALLDYPILOKKNSTGETLITSSYNGNYLAKFLVIRIGIKTSVLKNDCKYPLDIAL

AKGNTNIVCDLMKAGYN"

2025..4013

/gene="RP227"

2025..4013

/gene="RP227"

/codon\_start=1

/transl\_table=1

/product="DNA GYRASE SUBUNIT B (gyrB1)"

/protein\_id="CAA14690.1"

/db\_xref="GI:3860790"

/translation="MSDLFSLNKKEKNKIVTVNSAKDIEVLDGLEVRKRPQGMVIGG

TDSNAMHLYSEVLDNADMEAVAGFASIIIMKHQHSITIFDNGRGIPIDNHPKFPD

KSALIEVILTLHSGSKFSNNYHTSGGLHGVISVYNALSKHFKIKYKQKGLYSQSY

SKGAKLTDLISAEASKRLRGTSINFTPDPEIFSEKLFHFNPKKIYEIARSKAYLRGVS

IEWECEVEVPSDIPKALINPNGLKDYLSKISLDNLVPIEIPFSGNIESTVDDIKLE

WAICQNNDTSAFMQSYCNTVPTPGGTHEQGLKSAIILGLKAYSEMIGNKKAANLTI

EDILETASIVLSIFIVPSFOGOTKEKLVNSGVSKLVENIIKDHFDFLSSDKVLATH

LLEHVIAIAFERNNKNERNISRSKVTQKRLPGKLADCTFTSAEGTELEFVEGDSAG

GSQAQRNETQAVLPWLVKLVNVAASSTLEKIIINNAQDLEIALACGSLKANYKENL

RYEKIIITTDADVDGAHIALSLMTFFFLRMPKEGHLYLAKPPLVRLTOSNKIYYA

CDEEEKIKLYKLSKASKAKIEVGRFKGLGEMPAQLKTTMHPKASLLKVTLEDVQ

NVDKIVDLDLGMCKPEKRFQIYEQALVKMDQIINKLDI"

4122..5447

/gene="RP228"

4122..5447

/gene="RP228"

/codon\_start=1

/transl\_table=1

/product="TAIL-SPECIFIC PROTEASE PRECURSOR (ctp)"

/protein\_id="CAA14691.1"

/db\_xref="GI:3860791"

/translation="MLLRFIALLFSINCAIEGKETENKISNQEAYKQFQVFBRIEK

DYVQVPRQRMIDAINGLMNSLDPHNSYTTDELEDIFTFKGFGGIGVEIMYDSG

AIKIISSIDDLPTREIKVAGKIGDYIVGVNDELVSITLGNKAIKEMRGTPGTKVRLLIK

EEAKPOBELTPTREIKVPIKAHLEKNNIAYIRITTFNESTISELKAAYKVKLTESK

DNLKGILLDENNAGGILDOAIAVSDYFIDSGVITVTGRTTSSNSETKANEFSLKAP

KVPMIVLINGSASASIVAGALQDHKRAILLTGKTSFGKSGVQALQINSRAAVKLT

SKYITPSGRSIQAGEIPDILIEPAKVEIPEVKKIDKRFSESSLKNTLKNNAKKNL



coll and Saccharomyces cerevisiae and identification of an active site residue  
J. Biol. Chem. 271 (10), 5884-5891 (1996)  
MEDLINE 96215058  
REFERENCE 2 (bases 1 to 1078)  
AUTHORS Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S.D. and Schwab, H.  
TITLE Direct Submission  
JOURNAL Submitted (09-NOV-1995) Meinhard Hasslacher, Department of Biochemistry, University of Technology Graz, Petersgasse 12/2, Graz, 8010, Austria  
FEATURES  
source Location/Qualifiers  
1..1078  
/organism="Hevea brasiliensis"  
/db\_xref="taxon:3981"  
/tissue\_type="leaves"  
44..817  
/gene="hnl"  
44..817  
/gene="hnl"  
/codon\_start=1  
/product="hydroxynitrile lyase"  
/protein\_id="AAC49184.1"  
/db\_xref="gi:1223884"  
/translation="MAFAHFVLIHTICHGAWIHKLPLEALGHKVTALDLAASGVD  
PROTEIGSPDEYSEPLTFLAEPGEKVLVGGSGGLNATAADKYCEKIAAAVF  
HNSVLPDTEHCPSYVDKLMVEFPWDKDTVTYTKDGKEITGLKGLTLRLNLYL  
CGPEYELAKMLTRKSLFQNLAKRFFTEKYGSGSKYIYVTDQDEIFLPERQLWQ  
IENYKPKYVYVEGDHKLQTKTEIAELIQEVADTYN"  
BASE COUNT 324 a 191 c 239 g 324 t  
ORIGIN  
alignment\_scores  
Quality: 45.00 Length: 15  
Ratio: 3.750 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 66.667  
alignment\_block  
US-09-485-529-104 x HBU40402 ..  
Align seg 1/1 to: HBU40402 from: 1 to: 1078  
3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
||||| |||||||:||||| ||| |||:||||  
116 CTCTTGAGGCACTTGGCCACAGGTACTGCACTGGACCTTGCA 160  
seq\_name: gb\_pat:A59586  
seq\_documentation\_block: 1091 bp DNA PAT 06-MAR-1998  
LOCUS A59586  
DEFINITION Sequence 1 from Patent WO9703204.  
ACCESSION A59586  
VERSION A59586.1 GI:3714897  
KEYWORDS Para rubber tree.  
SOURCE Hevea brasiliensis  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Hevea.  
1 (bases 1 to 1091)  
AUTHORS Hasslacher, M., Schall, M., Schwab, H., Hayn, E.M., Kohlwein, S. and Griengl, H.  
TITLE (S)-HYDROXYNITRILYLASE FROM HEVEA BRASILIENSIS  
JOURNAL Patent: WO 9703204-A 1 30-JAN-1997;  
FEATURES  
source Location/Qualifiers  
1..1091  
/organism="Hevea brasiliensis"  
/db\_xref="taxon:3981"  
1..1091  
/gene="HNL"  
/function="(S)-HYDROXYNITRILYLASE"

gene 1..1091  
/gene="HNL"  
BASE COUNT 326 a 196 c 239 g 330 t  
ORIGIN  
alignment\_scores  
Quality: 45.00 Length: 15  
Ratio: 3.750 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 66.667  
alignment\_block  
US-09-485-529-104 x A59586 ..  
Align seg 1/1 to: A59586 from: 1 to: 1091  
3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
||||| |||||||:||||| ||| |||:||||  
129 CTCTTGAGGCACTTGGCCACAGGTACTGCACTGGACCTTGCA 173  
seq\_name: gb\_htg:AC026377  
seq\_documentation\_block: 161592 bp DNA HTG 06-NOV-2000  
LOCUS AC026377  
DEFINITION Mus musculus chromosome 11 clone RP23-267J8, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 64 unordered pieces.  
ACCESSION AC026377  
VERSION AC026377.4 GI:11094618  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 161592)  
AUTHORS Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, F., Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J., Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R., Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Martin, R., Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Oguh, M., Parish, B., Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G., Worley, K. and Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 161592)  
Worley, K.C.  
Direct Submission  
Submitted (22-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
Baylon Nov 4, 2000 this sequence version replaced gi:8248608.  
----- Genome Center  
Center: Baylor College of Medicine  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: MACL  
Center clone name: RP23-267J8  
----- Summary Statistics  
Sequencing vector: M13; L08821  
Chemistry: Dye-primer Bodipy: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 106250 bases at least Q40  
Consensus quality: 135137 bases at least Q30  
Consensus quality: 147325 bases at least Q20



\* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.



```
seq_documentation_block:
LOCUS AC020887 271519 bp DNA HTG 29-MAR-2000
DEFINITION Mus musculus clone RP23-472H3, WORKING DRAFT SEQUENCE, 86 unordered
pieces.
ACCESSION AC020887
VERSION AC020887.2 GI:7340304
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 271519)
DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 271519)
DOE Joint Genome Institute.
DIRECT SUBMISSION
SUBMITTED (10-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 29, 2000 this sequence version replaced gi:6686421.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 168487 bases at least Q40
Consensus quality: 217951 bases at least Q30
Consensus quality: 233703 bases at least Q20
Estimated insert size: 271519; sum-of-contigs estimation
Estimated insert size: 205000; pulse field gel estimation
Quality coverage: 3.76x in Q20 bases; pulse field gel estimation
Quality coverage: 2.84x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 86 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1034: contig of 1034 bp in length
* gap of unknown length
* 1035 2087: contig of 1053 bp in length
* gap of unknown length
* 2088 3169: contig of 1082 bp in length
* gap of unknown length
* 3170 4223: contig of 1054 bp in length
* gap of unknown length
* 4224 5316: contig of 1093 bp in length
* gap of unknown length
* 5317 6505: contig of 1189 bp in length
* gap of unknown length
* 6506 7563: contig of 1058 bp in length
* gap of unknown length
* 7564 8823: contig of 1260 bp in length
* gap of unknown length
* 8824 9948: contig of 1125 bp in length
* gap of unknown length
* 9949 11246: contig of 1298 bp in length
* gap of unknown length
* 11247 12253: contig of 1007 bp in length
* gap of unknown length
* 12254 13595: contig of 1342 bp in length
* gap of unknown length
* 13596 14629: contig of 1034 bp in length
* gap of unknown length
* 14630 15690: contig of 1061 bp in length
* gap of unknown length
* 15691 16855: contig of 1165 bp in length
* gap of unknown length
* 16856 18146: contig of 1291 bp in length
* gap of unknown length
* 18147 19341: contig of 1195 bp in length
* gap of unknown length
* 19342 21053: contig of 1710 bp in length
* gap of unknown length
* 21052 22054: contig of 1003 bp in length
* gap of unknown length
* 22055 23447: contig of 1393 bp in length
* gap of unknown length
* 23448 24766: contig of 1319 bp in length
* gap of unknown length
* 24767 26376: contig of 1610 bp in length
* gap of unknown length
* 26377 27634: contig of 1258 bp in length
* gap of unknown length
* 27635 28682: contig of 1048 bp in length
* gap of unknown length
* 28683 29852: contig of 1170 bp in length
* gap of unknown length
* 29853 31137: contig of 1285 bp in length
* gap of unknown length
* 31138 32646: contig of 1509 bp in length
* gap of unknown length
* 32647 33673: contig of 1027 bp in length
* gap of unknown length
* 33674 35005: contig of 1332 bp in length
* gap of unknown length
* 35006 36149: contig of 1144 bp in length
* gap of unknown length
* 36150 37211: contig of 1062 bp in length
* gap of unknown length
* 37212 38258: contig of 1047 bp in length
* gap of unknown length
* 38259 39463: contig of 1205 bp in length
* gap of unknown length
* 39464 40738: contig of 1275 bp in length
* gap of unknown length
* 40739 41915: contig of 1177 bp in length
* gap of unknown length
* 41916 42959: contig of 1044 bp in length
* gap of unknown length
* 42960 44203: contig of 1244 bp in length
* gap of unknown length
* 44204 45244: contig of 1041 bp in length
* gap of unknown length
* 45245 46267: contig of 1023 bp in length
* gap of unknown length
* 46268 47576: contig of 1309 bp in length
* gap of unknown length
* 47577 48726: contig of 1150 bp in length
* gap of unknown length
* 48727 49889: contig of 1163 bp in length
* gap of unknown length
* 49890 50897: contig of 1008 bp in length
* gap of unknown length
* 50898 52130: contig of 1233 bp in length
* gap of unknown length
* 52131 53594: contig of 1464 bp in length
* gap of unknown length
* 53595 54946: contig of 1352 bp in length
* gap of unknown length
* 54947 56491: contig of 1545 bp in length
* gap of unknown length
* 56492 57665: contig of 1174 bp in length
* gap of unknown length
* 57666 58743: contig of 1078 bp in length
* gap of unknown length
* 58744 60730: contig of 1987 bp in length
* gap of unknown length
* 60731 62189: contig of 1459 bp in length
* gap of unknown length
* 62190 64258: contig of 2069 bp in length
```





```
TITLE      CHST1 and CHST2 sulfotransferases expressed by human vascular
localization
JOURNAL     Genomics 55 (3), 345-347 (1999)
MEDLINE     99168906
REFERENCE   2 (bases 1 to 1241)
AUTHORS     Li,X. and Tedder,T.F.
TITLE       Direct Submission
JOURNAL     Submitted (03-SEP-1998) Immunology, Duke University Medical Center,
Durham, NC 27710, USA
FEATURES
  source    1..1241
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="umbilical vein endothelium"
  gene      1..1241
            /gene="CHST1"
  CDS       5..1240
            /gene="CHST1"
            /codon_start=1
            /product="carbohydrate sulfotransferase 1"
            /protein_id="AA019878.1"
            /db_xref="GI:4406152"
            /translation="MQCSKAVLLALLASTAIQYTAIRTFKSFHTCPGLAERGLAE
RLCEESTFYAINSRKTHILILATFRSGSVFGQLFNQHLDDVFLFPLHYHVQNTLIP
RFTQKSPADRRVMLGASRDLLRSLYDCDLYFLENYIKPPPVNHTTDRIFRGRASRYL
CSRVCPPGPPADLVLEEGDCVRKGLNLTVAACRKRSHVAIKTVRVVNDLRA
LVEDPRNLKVIQIVLRDPGRGILASRSETDRTYRLRLWXTGTRKPYNLDTQLTVC
EDFSNYSVTGLMRPPLKGYMLVRYEDLARLNPMLKTEEYIGFLGIPLDHSVARWION
NTRGDPTLGHKHYGTVRNSATAKWRFLRSYDIVAFQACQVLAQLGKYKIAASEE
ELKNPVSLSVEERDFRPF5"
BASE COUNT 222 a 459 c 363 g 197 t
ORIGIN

alignment_scores:
  Quality: 44.00      Length: 14
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 85.714  Percent Identity: 57.143

alignment_block:
US-09-485-529-104 x AF090137 ..
Align seg 1/1 to: AF090137 from: 1 to: 1241
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1136 CAGGTGCTGCGCCAGCTGGCTACAAGATCGCGCTCGGAG 1177

seq_name: gb_pat:E16306
seq_documentation_block:
LOCUS      E16306      1458 bp      DNA      PAT      28-JUL-1999
DEFINITION CDNA encoding keratan sulphate 6-sulfo-transferase.
ACCESSION  E16306
VERSION    E16306.1 GI:5710989
KEYWORDS   Homo sapiens.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1458)
AUTHORS   Fukuda,M. and Hanebuch,N.
TITLE     KERATAN SULFATE 6-SULFOTRANSFERASE AND DNA CODING FOR THE SAME
JOURNAL   Patent: JP 1998155488-A 1 16-JUN-1998;
          SEIKAGAKU KOGYO CO LTD
COMMENT   OS Homo sapiens (human)
          PN JP 1998155488-A/1
          PD 16-JUN-1998
          PF 29-NOV-1996 JP 1996320535
          PI FUKUDA MASAKAZU, HANEBUCHI NAGAMOTO
          PC C12N15/09,C07H21/04,C07K14/47,C08B37/00,C12N9/10, PC
          C12P19/26/C12P21/02,
```

```
PC      (C12N15/09,C12R1:91),(C12N9/10,C12R1:19);
CC      strandedness: Single;
CC      topology: linear;
CC      hypothetical: No;
CC      anti-sense: No;
FH      Key
FH      Location/Qualifiers
FT      source    1..1458
            /organism="Homo sapiens"
            /tissue_type="Brain"
FT      CDS       1..1236
            /product="keratan sulphate
            6-sulfo-transferase"
            1237..1458.
FT      3'UTR     Location/Qualifiers
            1..1458
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 254 a 526 c 421 g 257 t
ORIGIN

alignment_scores:
  Quality: 44.00      Length: 14
  Ratio: 3.667       Gaps: 0
  Percent Similarity: 85.714  Percent Identity: 57.143

alignment_block:
US-09-485-529-104 x E16306 ..
Align seg 1/1 to: E16306 from: 1 to: 1458
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1132 CAGGTGCTGCGCCAGCTGGCTACAAGATCGCGCTCGGAG 1173

seq_name: gb_ba:BSORFRM1
seq_documentation_block:
LOCUS      BSORFRM1    2096 bp      DNA      BCT      27-JAN-1997
DEFINITION B.subtilis orfRM1 DNA.
ACCESSION  X98341
VERSION    X98341.1 GI:1402943
KEYWORDS   orfRM1 DNA.
SOURCE     Bacillus subtilis.
ORGANISM   Bacillus subtilis
REFERENCE  1 (bases 1 to 2096)
AUTHORS   Marasco,R., Varcamonti,M., Ricca,E. and Sacco,M.
TITLE     A new Bacillus subtilis gene with homology to Escherichia coli prc
JOURNAL   Gene 183 (1-2), 149-152 (1996)
MEDLINE   97149292
REFERENCE  2 (bases 1 to 2096)
AUTHORS   Sacco,M.
TITLE     Direct Submission
JOURNAL   Submitted (06-JUN-1996) M. Sacco, Istituto Internazionale di
          Genetica e Biofisica, via Marconi 10, 80125, Napoli ITALY
FEATURES
  source    1..2096
            /organism="Bacillus subtilis"
            /strain="PY17"
            /db_xref="taxon:1423"
            535..540
            558..563
            570..1992
            592..1992
            /gene="orfRM1"
            592..1992
            /gene="orfRM1"
            /codon_start=1
            /transl_table=11
            /protein_id="CAA66987.1"
            -35_signal
            -10_signal
            mRNA
            gene
            CDS
```

```

/db_xref="GI:1402944"
/db_xref="SPTREMBL:Q45645"
/translation="MKRQLKLEFVILITAVVASALTLFITGNSSILGQKSASTGDSKP
DLNKAAYEQIKSDYQKTDKDDKGLGAIKGMIOQSDDPYSTYMDQEQAKSFDETISAS
FGIGAEVEEKDEGELIVSPIKSPAEKAGIKRPOQIIKVNKSGKGMNVNEAVALIR
GGKGVKLELNAGVGNIDLSIKRDTIPVETVYSEMKNNGICEIQIITSFSETAKEL
TDAIDSLRKGAKGVILDRGNPGGLMEQALMSNLFIDKKNIMQVEYKNGSKEYMK
AEKRVKTPVTVLVNDGTASRNVGRCLHSSNVPLIGETTFGKGTVTOTAKKEYDDG
SVTKLTAKWLADGEWHKKGKIPQVKAELPDYAKLPYLDADKTYKSGDTCNNKYA
QMKLALGFKYKVNMSYDQDVSFVVKQFOKKELNETGLTGTITTKLMIELQKKLSD
NDTQMEKALETUKKEM"
terminator 711 a 409 c 482 g 494 t
BASE COUNT 711 a 409 c 482 g 494 t
ORIGIN

alignment_scores:
  Quality: 44.00      Length: 15
  Ratio: 3.143      Gaps: 0
Percent Similarity: 93.333 Percent Identity: 53.333

alignment_block:
US-09-485-529-104 x BSORFRM1 ..
Align seg 1/1 to: BSORFRM1 from: 1 to: 2096

2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMet 16
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1771 AAAAGTCTCAAGCAGCTGGTGTACAGGTAAGGTTACAGATATG 1815

seq_name: gb_pr:HSU65637

seq documentation_block:
LOCUS HSU65637 2190 bp mRNA PRI 06-AUG-1998
DEFINITION Homo sapiens chondroitin-6-sulfotransferase mRNA, complete cds.
ACCESSION U65637
VERSION U65637.1 GI:2769701
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2190)
AUTHORS Williams,K.J.
TITLE Atherosclerosis: cell biology and lipoproteins
JOURNAL Curr. Opin. Lipidol. 7 (6), U202-U208 (1996)
MEDLINE 97189336
REFERENCE 2 (bases 1 to 2190)
AUTHORS Mazany,K.D., Peng,T., Watson,C.E., Tabas,I. and Williams,K.J.
TITLE Human chondroitin 6-sulfotransferase: cloning, gene structure, and
chromosomal localization
JOURNAL Biochim. Biophys. Acta 1407 (1), 92-97 (1998)
MEDLINE 98306085
REFERENCE 3 (bases 1 to 2190)
AUTHORS Peng,T., Tabas,I. and Williams,K.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1996) Medicine, Thomas Jefferson University, 1020
Locust Street, Philadelphia, PA 19107-6799, USA
FEATURES
  source
    1. .2190
      Location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /note="part of the sequence corresponds to an EST, GenBank
        chondroitin-6-sulfotransferase cDNA sequence, GenBank
        Accession Number D49915"
        125. .1360
        /codon_start=1
        /product="chondroitin-6-sulfotransferase"
        /protein_id="AAC28776.1"
        /db_xref="GI:2769702"
        /translation="MQCSKAVLLALASIAIOYTAIRTFKSFHTCPLAEAGLAE
        RLCEESPTAYNLSRTHILATRRSGSFVQLFNQHLVDVLYFELYHVQNTLIP
        CDS
    1. .1360
      Location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /note="part of the sequence corresponds to an EST, GenBank
        chondroitin-6-sulfotransferase cDNA sequence, GenBank
        Accession Number D49915"
        125. .1360
        /codon_start=1
        /product="chondroitin-6-sulfotransferase"
        /protein_id="AAC28776.1"
        /db_xref="GI:2769702"
        /translation="MQCSKAVLLALASIAIOYTAIRTFKSFHTCPLAEAGLAE
        RLCEESPTAYNLSRTHILATRRSGSFVQLFNQHLVDVLYFELYHVQNTLIP

```

```

RFTQKSPADRRVLMGASRDLLSLYDCDLYFLENYIKPPVNHHTDRIFFRGSARVL
CSRVCDDPGPADLVLEEGDCVRKCGLLNLTVAAEACRERSHVAIKTVRVPENDLRA
LYEDPRLNLKVIQLVRDPGCLLASRSETPRDTYRLRLMWGTGRKPYNLDTVQLTVC
EDFSNSVSTGLMRPWLKGYMLVRYEDLARPNMKTEEYIGFLGIPLDSSHVARWTON
NTRGDPDLGKHKGTVNRNSAATAEKWRFLSYDIVAFQAQACQOVLQALGYKTAASEE
ELKNPSVSLVEERDFRPF"
BASE COUNT 429 a 717 c 595 g 449 t
ORIGIN

alignment_scores:
  Quality: 44.00      Length: 14
  Ratio: 3.667      Gaps: 0
Percent Similarity: 85.714 Percent Identity: 57.143

alignment_block:
US-09-485-529-104 x HSU65637 ..
Align seg 1/1 to: HSU65637 from: 1 to: 2190

2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1256 CAGGTGCTGGCCAGCTGGCTACAGATCGCCGCTCGGAG 1297

seq_name: gb_pr:AB003791

seq documentation_block:
LOCUS AB003791 2415 bp mRNA PRI 14-FEB-1998
DEFINITION Homo sapiens mRNA for keratan sulfate Gal-6-sulfotransferase,
complete cds.
ACCESSION AB003791
VERSION AB003791.1 GI:2887402
KEYWORDS keratan sulfate Gal-6-sulfotransferase.
SOURCE Homo sapiens fetus brain cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2415)
AUTHORS Habuchi,O.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1997) to the DBJ/EMBL/GenBank databases. Osami
Habuchi, Aichi University of Education, Department of Life Science;
Igaya-cho, Kariya, Aichi 448, Japan (Tel:0556-36-3111,
Fax:0556-36-4337)
REFERENCE 2 (sites)
AUTHORS Fukuta,M., Inazawa,J., Torii,T., Tsuzuki,K., Shimada,E. and
Habuchi,O.
TITLE Molecular cloning and characterization of human keratan sulfate
Gal-6-sulfotransferase
JOURNAL J. Biol. Chem. 272 (51), 32321-32328 (1997)
MEDLINE 98070405
FEATURES
  Location/Qualifiers
    1. .2415
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="11"
      /dev_stage="fetus"
      /map="11p11.1-11.2"
      /tissue_type="brain"
      367. .1602
      /function="enzyme"
      /codon_start=1
      /product="keratan sulfate Gal-6-sulfotransferase"
      /protein_id="BAA24840.1"
      /db_xref="GI:2887403"
      /translation="MQCSKAVLLALASIAIOYTAIRTFKSFHTCPLAEAGLAE
      RLCEESPTAYNLSRTHILATRRSGSFVQLFNQHLVDVLYFELYHVQNTLIP
      RFTQKSPADRRVLMGASRDLLSLYDCDLYFLENYIKPPVNHHTDRIFFRGSARVL
      CSRVCDDPGPADLVLEEGDCVRKCGLLNLTVAAEACRERSHVAIKTVRVPENDLRA
      LYEDPRLNLKVIQLVRDPGCLLASRSETPRDTYRLRLMWGTGRKPYNLDTVQLTVC
      EDFNSVSTGLMRPWLKGYMLVRYEDLARPNMKTEEYIGFLGIPLDSSHVARWTON
      NTRGDPDLGKHKGTVNRNSAATAEKWRFLSYDIVAFQAQACQOVLQALGYKTAASEE
      ELKNPSVSLVEERDFRPF"

```

```

BASE COUNT      437 a      805 c      687 g      486 t
ORIGIN
alignment_scores:
  Quality:      44.00      Length:      14
  Ratio:        3.667      Gaps:        0
  Percent Similarity: 85.714      Percent Identity: 57.143
alignment_block:
US-09-485-529-104 x AB003791 ..
Align seg 1/1 to: AB003791 from: 1 to: 2415
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::..:||||| |||||::: |||||:::
1498 CAGGTGCTGCCAGCTGGGCTACAGATCGCGCCCTCGGAG 1539
seq_name: gb_v1:AF298771
seq_documentation_block:
LOCUS AF298771 4374 bp RNA VRL 09-OCT-2000
DEFINITION Porcine reproductive and respiratory syndrome virus RNA-dependent
RNA polymerase gene, partial cds.
ACCESSION AF298771
VERSION AF298771.1 GI:10719655
KEYWORDS
SOURCE Porcine reproductive and respiratory syndrome virus.
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Arteriviridae; Arterivirus.
REFERENCE
AUTHORS Bautista,E.M. and McGruder,E.D.
TITLE Cloning and expression of PRRSV ORF1b and demonstration and
characterization of its NTPase activity
JOURNAL Unpublished
REFERENCE
AUTHORS Bautista,E.M. and McGruder,E.D.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Biology Research and Development, Elanco
Animal Health, a Division of Eli Lilly and Company, 2001 West Main
St, Greenfield, IN 46140, USA
FEATURES
source
/organism="Porcine reproductive and respiratory syndrome
virus"
/isolate="PRRSV14"
/db_xref="taxon:28344"
<1..4374
/note="ORF1b; zinc-finger protein; helicase (NTPase);
expression of ORF1b occurs due to an upstream ribosomal
frameshift"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="AAG22078.1"
/db_xref="GI:10719656"
/translation="LAASGLRRCGRGLVVTETAVKIVKFNHRTFTLGPVNLKVASEV
ELKDVRHNQHPVAPRTDGGVLLRSVPSLIDVLISGADASKLLAHGPGNTGIDG
TLWDFESPATKEEVALSQIIOACDIRGDAPRGLPKLYVRGNPERVKGVLQNTF
FGDIPYTPSTGSPVHAACLTNAPVTDGRSVLATTPPGFELYVPTIPASVLDY
LDSRPDCPKTQHEGCEADALKDLSKYDLSQGFVLPGLVRLVRKRYLFAHVKGCPVHV
RPSTYPAKNSMAGINRFRPTKIQSVPEIDVLCQAQVRENWQTVPTCLKKQYCGKK
KTRTILGTNNFIALAHNRLSGVTQGGPKAFNSPIALGNKNKFELOTPVLGRGLEAD
LASCDRSTPAIWRPFAANLLVELACAEHLPVYVNLCCDILLVTSQGAVTKRGGISSG
DPITVSNTIYSLVIAQHNVLSYKESHPHGLLFLODLKFEEDMLKQVPLIVSDLI
VLYAESPTMPNHNWVHLNMLGFDQDPKKATDTSFSLGCRINGRLVPMNRDRI
LAALAYHMKANVSEIYASAAAIIIMDCACLEYDPEWFEELVGVIAQCCARKDGSFPG
TLPFMSMWEKLSRNEKRGKRCVCGAPATACGLDVCYIYTHFHQCPYIWCQ
HPAGSGSCSECKSPVKGTSPLDVEQVYKPPRTVIMHVEQGLTPLDPRGYQTRRG
LVSVRRGIRGNEVELPDGYASTALLPQTEIKNMVASNVLRSRFTIIGPPGACKTYW
LQQQVQDGVITYTTHOTMLDMIRALGTCRFNPAGTTLQFPVPSRTGPPWVRIAGGW
CPGNSTLDEAAYCNHLDVLRLLSKTLTLCGLDQKQLHPVGFDSHCYFVDIMPOTQLK
TIWRFGQNICDAIQDYNDKLSMWNTTTRVTHVEKPVIRGVQVLTPIYHHRDEDDAITID
SSGATEDVVTLLHLPTKDSLNRQALVAITRARHAIIFYYDPHROLQGFDFDLPKAGTPV
NLAVHRDQQLIVLDNRNKECTVAQALNGDKFRATDKRVVDSLRALCADLEGSSPLP
KVAHNLGFYFSPDLTQFAKLPVLAHPVVTQNNKEKWPDLRLVLRPIHKYSRACI
GAGYMGVSPFLGTGGVSYLTTFKVGEOQLPETVSTGRIEVDCREYLDREYR
AASLPFAFIGDVKGCTTGGCHHTVSRVLPVLPKESVAVVGVSSPGKAALCTLTDV
YLPDLERAYLHPETOSKCKWMMLDKFEVRLVWVKDXTAYFQLEGRYFTWYQLASVSYI
RVPVNSTVYLDPCMGPALCNRVVGSTHMGADLAVTPYDGAKIILSSAYHGEPPGY
KILACAEFSLDDPVYKHTWGFESDTAYLYEFTGTGDEWDYNDAFRAREQEGTYKAT
ATSLKFYFPPGPVIEPTILGN"
BASE COUNT      1029 a      1202 c      1104 g      1039 t
ORIGIN
alignment_scores:
  Quality:      44.00      Length:      17
  Ratio:        2.750      Gaps:        0
  Percent Similarity: 94.118      Percent Identity: 47.059
alignment_block:
US-09-485-529-104 x AF298771 ..
Align seg 1/1 to: AF298771 from: 1 to: 4374
1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1690 GACAGGATCCTCGCGCGCCTCGCCTATCACATGAGCGGAGTAATGTTTC 1739
17 a 17
1740 T 1740
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2001, 17:03:44 ; Search time 23.9 Seconds  
(without alignments)  
52.688 Million cell updates/sec

Title: US-09-485-529-104

Perfect score: 81

Sequence: 1 DELAALGVKVRSDMA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1:	/SID88/gcgdata/genseq/genseq/AA1980.DAT.*
2:	/SID88/gcgdata/genseq/genseq/AA1981.DAT.*
3:	/SID88/gcgdata/genseq/genseq/AA1982.DAT.*
4:	/SID88/gcgdata/genseq/genseq/AA1983.DAT.*
5:	/SID88/gcgdata/genseq/genseq/AA1984.DAT.*
6:	/SID88/gcgdata/genseq/genseq/AA1985.DAT.*
7:	/SID88/gcgdata/genseq/genseq/AA1986.DAT.*
8:	/SID88/gcgdata/genseq/genseq/AA1987.DAT.*
9:	/SID88/gcgdata/genseq/genseq/AA1988.DAT.*
10:	/SID88/gcgdata/genseq/genseq/AA1989.DAT.*
11:	/SID88/gcgdata/genseq/genseq/AA1990.DAT.*
12:	/SID88/gcgdata/genseq/genseq/AA1991.DAT.*
13:	/SID88/gcgdata/genseq/genseq/AA1992.DAT.*
14:	/SID88/gcgdata/genseq/genseq/AA1993.DAT.*
15:	/SID88/gcgdata/genseq/genseq/AA1994.DAT.*
16:	/SID88/gcgdata/genseq/genseq/AA1995.DAT.*
17:	/SID88/gcgdata/genseq/genseq/AA1996.DAT.*
18:	/SID88/gcgdata/genseq/genseq/AA1997.DAT.*
19:	/SID88/gcgdata/genseq/genseq/AA1998.DAT.*
20:	/SID88/gcgdata/genseq/genseq/AA1999.DAT.*
21:	/SID88/gcgdata/genseq/genseq/AA2000.DAT.*
22:	/SID88/gcgdata/genseq/genseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	17	20	AA02537 Peptide derived fr
2	81	100.0	623	20	AA02540 Protein encoded by
3	78	96.3	100	20	AA02542 Protein encoded by
4	78	96.3	123	20	AA02543 Protein encoded by
5	78	96.3	256	20	AA02538 Protein encoded by
6	78	96.3	630	20	AA02541 Protein encoded by
7	70	86.4	517	21	AA02547 Arabidopsis thalia
8	70	86.4	518	21	AA02548 Arabidopsis thalia
9	70	86.4	531	21	AA02549 Arabidopsis thalia
10	70	86.4	532	18	AA02545 Arabidopsis thalia
11	70	86.4	532	22	AA02545 A. thaliana transc

12	70	86.4	532	22	AAE01907 Arabidopsis thalia
13	70	86.4	533	21	AAG38575 Arabidopsis thalia
14	70	86.4	587	21	AAB28574 Arabidopsis thalia
15	70	86.4	587	22	AAE02560 A. thaliana transc
16	70	86.4	587	22	AAE01892 Arabidopsis thalia
17	66	81.5	572	22	AAE01883 Amino acid sequenc
18	66	81.5	572	22	AAE01884 Amino acid sequenc
19	62	76.5	138	20	AA02544 Hevea brasiliensis
20	45	55.6	257	18	AAW29164 Hevea brasiliensis
21	45	55.6	257	18	AAW10593 Hevea brasiliensis
22	44	54.3	411	19	AAW61100 Keratan sulphate 6
23	44	54.3	1457	22	AAE02566 Porcine reproducti
24	42	51.9	130	21	AAG44380 Arabidopsis thalia
25	42	51.9	134	21	AAG44379 Arabidopsis thalia
26	42	51.9	263	21	AAG08858 Arabidopsis thalia
27	42	51.9	263	21	AAG43311 Arabidopsis thalia
28	42	51.9	269	21	AAG08857 Arabidopsis thalia
29	42	51.9	269	21	AAG43310 Arabidopsis thalia
30	41	50.6	77	15	AAE02586 CBEV spheroidin.
31	41	50.6	77	15	AAE02587 CBEV spheroidin.
32	41	50.6	77	20	AA030179 Protein encoded by
33	41	50.6	77	20	AA030180 Protein encoded by
34	41	50.6	491	22	AAG98407 Escherichia coli p
35	41	50.6	506	21	AAE02587 Neisseria gonorrhe
36	41	50.6	506	21	AAE02587 Neisseria meningit
37	41	50.6	506	21	AAE02587 Neisseria meningit
38	40	49.4	200	22	AAG82046 S. epidermidis ope
39	40	49.4	200	22	AAG82048 S. epidermidis ope
40	40	49.4	487	22	AAU01923 Pseudomonas aerugi
41	39	48.1	140	21	AAG32315 Arabidopsis thalia
42	39	48.1	263	21	AAG29564 Arabidopsis thalia
43	39	48.1	271	21	AAG29563 Arabidopsis thalia
44	39	48.1	398	22	AAW40727 Human polypeptide
45	39	48.1	550	22	AAU03810 G protein-coupled

#### ALIGNMENTS

RESULT 1  
AA02537  
ID AAY02537 standard; Peptide; 17 AA.  
XX  
AC AAY02537;  
XX  
DT 16-JUL-1999 (first entry)  
XX  
DE Peptide derived from a protein causing growth inhibition.  
DE Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol.  
XX  
XX Unidentified.  
XX OS  
XX PN WO9909174-A1.  
XX  
XX PD 25-FEB-1999.  
XX  
XX PF 07-AUG-1998; 98WO-GB02383.  
XX  
XX PR 13-AUG-1997; 97GB-0017192.  
XX  
XX PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
XX PI Harberd NP, Peng J, Richards DE;  
XX  
XX DR WPI; 1999-181040/15.  
XX  
XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
XX PT provides inhibition of the growth of plants, which inhibition is  
XX PT antagonised by gibberellin, used to confer a dwarf phenotype



PS Claim 1; Page 53; 88pp; English.

XX The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in *Triticum aestivum*, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence represents  
CC a peptide derived from a protein causing growth inhibition when  
CC expressed in wheat.

SQ Sequence 17 AA;

Query Match 100.0%; Score 81; DB 20; Length 17;

Best Local Similarity 100.0%; Pred. No. 1.8e-08; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

Qy 1 DELLAALGKVKVRSMDMA 17  
Db 1 dellaalgkvrasm 17

RESULT 2

RAY02540  
ID AAY02540 standard; Protein; 623 AA.

XX AC AAY02540;

XX DT 16-JUL-1999 (first entry)

XX DE Protein encoded by wheat Rht clone 5a1 genomic sequence.

XX KW Rht gene; homologue; *Triticum aestivum*; wheat; growth inhibition;  
XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
XX KW paclobutrazol.

XX OS *Triticum aestivum*.

XX PN WO9909174-A1.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-GB02383.

XX PR 13-AUG-1997; 97GB-0017192.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Harberd NP, Peng J, Richards DE;

XX DR WPI; 1999-181040/15.

XX DR N-PSDB; AAX36279.

XX PT New *Triticum aestivum* polynucleotides - encode a polypeptide which  
XX PT provides inhibition of the growth of plants, which inhibition is  
XX PT antagonised by gibberellin, used to confer a dwarf phenotype

XX PS Disclosure; Fig 8b; 88pp; English.

XX The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in *Triticum aestivum*, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be

CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence is encoded by  
CC the wheat Rht clone 5a1 genomic sequence.

SQ Sequence 623 AA;

Query Match 100.0%; Score 81; DB 20; Length 623;

Best Local Similarity 100.0%; Pred. No. 1e-06; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

Qy 1 DELLAALGKVKVRSMDMA 17  
Db 38 dellaalgkvrasm 54

RESULT 3

RAY02542  
ID AAY02542 standard; Protein; 100 AA.

XX AC AAY02542;

XX DT 16-JUL-1999 (first entry)

XX DE Protein encoded by the partial sequence of the maize D8-1 allele.

XX KW Rht gene; homologue; *Triticum aestivum*; wheat; growth inhibition;  
XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
XX KW paclobutrazol; maize.

XX OS Zea mays.

XX PN WO9909174-A1.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-GB02383.

XX PR 13-AUG-1997; 97GB-0017192.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Harberd NP, Peng J, Richards DE;

XX DR WPI; 1999-181040/15.

XX DR N-PSDB; AAX36281.

XX PT New *Triticum aestivum* polynucleotides - encode a polypeptide which  
XX PT provides inhibition of the growth of plants, which inhibition is  
XX PT antagonised by gibberellin, used to confer a dwarf phenotype

XX PS Disclosure; Fig 11b; 88pp; English.

XX The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in *Triticum aestivum*, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence is encoded by  
CC the partial sequence of the maize D8-1 allele.

```

SQ Sequence 100 AA;
Query Match 96.3%; Score 78; DB 20; Length 100;
Best Local Similarity 94.1%; Pred. No. 4.7e-07;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17
|||||
Db 34 dellaalgykvrssdma 50

RESULT 4
AA02543
ID AAY02543 standard; Protein; 123 AA.
XX AC AAY02543;
XX DT 16-JUL-1999 (first entry)
XX DE Protein encoded by the partial sequence of the maize D8-2023 allele.
XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;
XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;
XX KW paclobutrazol; maize.
XX OS Zea mays.
XX PN WO9509174-A1.
XX PD 25-FEB-1999.
XX PF 07-AUG-1998; 98WO-GB02383.
XX PR 13-AUG-1997; 97GB-0017192.
XX PA (PLAN-) PLANT BIOSCIENCE LTD.
XX PI Harberd NP, Peng J, Richards DE;
XX DR N-PSDB; AAX36282.
XX DR WPI: 1999-181040/15.
XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which
XX PT provides inhibition of the growth of plants, which inhibition is
XX PT antagonised by gibberellin, used to confer a dwarf phenotype
XX PS Disclosure; Fig 1ld; 88pp; English.
XX CC The specification describes polypeptides encoded by the Rht gene (and
XX CC its homologues) that, when expressed in Triticum Aestivum, inhibit
XX CC growth of the plant. This growth inhibition is antagonised by
XX CC gibberellin. The products can be used to provide Rht expression in
XX CC plants, conferring a dwarf phenotype on a plant which is correctable
XX CC by treatment with gibberellin. In addition, the products can be
XX CC used to produce Rht mutant plants which are dwarfed compared with
XX CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants
XX CC may be made by knocking out Rht or the relevant homologous gene in
XX CC the plant of interest. Plants may be made which are resistant to
XX CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,
XX CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds
XX CC dwarf but let crop plants grow tall. The present sequence is encoded by
XX CC the partial sequence of the maize D8-2023 allele.
XX SQ Sequence 123 AA;

Query Match 96.3%; Score 78; DB 20; Length 123;
Best Local Similarity 94.1%; Pred. No. 6e-07;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17
|||||
Db 38 dellaalgykvrssdma 54

RESULT 6
AA02541
ID AAY02541 standard; Protein; 630 AA.
XX AC AAY02541;

```

---

```

Db 22 dellaalgykvrssdma 38

RESULT 5
AA02538
ID AAY02538 standard; Protein; 256 AA.
XX AC AAY02538;
XX DT 16-JUL-1999 (first entry)
XX DE Protein encoded by rice EST D39460 sequence.
XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;
XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;
XX KW paclobutrazol; rice; expressed sequence tag; EST.
XX OS Oryza sativa.
XX PN WO9909174-A1.
XX PD 25-FEB-1999.
XX PF 07-AUG-1998; 98WO-GB02383.
XX PR 13-AUG-1997; 97GB-0017192.
XX PA (PLAN-) PLANT BIOSCIENCE LTD.
XX PI Harberd NP, Peng J, Richards DE;
XX DR N-PSDB; AAX36277.
XX DR WPI: 1999-181040/15.
XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which
XX PT provides inhibition of the growth of plants, which inhibition is
XX PT antagonised by gibberellin, used to confer a dwarf phenotype
XX PS Claim 12; Fig 6b; 88pp; English.
XX CC The specification describes polypeptides encoded by the Rht gene (and
XX CC its homologues) that, when expressed in Triticum Aestivum, inhibit
XX CC growth of the plant. This growth inhibition is antagonised by
XX CC gibberellin. The products can be used to provide Rht expression in
XX CC plants, conferring a dwarf phenotype on a plant which is correctable
XX CC by treatment with gibberellin. In addition, the products can be
XX CC used to produce Rht mutant plants which are dwarfed compared with
XX CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants
XX CC may be made by knocking out Rht or the relevant homologous gene in
XX CC the plant of interest. Plants may be made which are resistant to
XX CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,
XX CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds
XX CC dwarf but let crop plants grow tall. The present sequence is encoded by
XX CC wheat Rht gene.
XX SQ Sequence 256 AA;

Query Match 96.3%; Score 78; DB 20; Length 256;
Best Local Similarity 94.1%; Pred. No. 1.4e-06;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17
|||||
Db 38 dellaalgykvrssdma 54

RESULT 6
AA02541
ID AAY02541 standard; Protein; 630 AA.
XX AC AAY02541;

```



PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144634.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145319.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 86.4%; Score 70; DB 21; Length 517;  
Best Local Similarity 82.4%; Pred. No. 9.4e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
|||||:|||||:  
Db 12 dellavlgkyvrssema 28

## RESULT 8

AAG38576  
ID AAG38576 standard; Protein; 518 AA.

XX AC AAG38576;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47611.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125789.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.

```

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.4%; Score 70; DB 21; Length 518;
Best Local Similarity 82.4%; Pred. No. 9.5e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17
Db 13 dellavlgkvrssma 29
||||| |||||:|:|

RESULT 9
AAB28575
ID AAB28575 standard; Protein; 531 AA.
XX
AC AAB28575;
XX
DT 09-FEB-2001 (first entry)
XX
DE Arabidopsis SCLa2.
XX
KW Scarecrow gene; SCR gene; plant; regulatory gene; SCARECROW-like; SCL;
KW transgenic plant; cell division; molecular marker; herbicide resistance;
KW salt resistance; pathogen resistance; insect resistance.
XX
OS Arabidopsis thaliana.
XX
PN WO200053723-A2.
XX
PD 14-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-US05875.
XX
PR 10-MAR-1999; 99US-0265585.
XX
PA (UYNV ) UNIV NEW YORK STATE.
XX
PI Benfey PN, Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysch L;
PI Helariutta Y, Bruce W, Lim J;

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.4%; Score 70; DB 21; Length 518;
Best Local Similarity 82.4%; Pred. No. 9.5e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17
Db 27 dellavlgkvrssma 43
||||| |||||:|:|

RESULT 10
AAB30792
ID AAB30792 standard; Protein; 532 AA.
XX
AC AAB30792;
XX
DT 19-MAR-1998 (first entry)
XX
DE Arabidopsis thaliana gibberellin insensitivity gai gene product.
XX
KW Gibberellin insensitivity; gai; plant growth inhibition;
KW dwarf phenotype; lodging resistance; increased yield;
KW flowering regulation; bolting inhibition; spinach; lettuce;
KW antibody; identification; probe; primer; antisense; sense;
KW expression regulation; co-suppression; rice;
KW Bakane disease resistance.
XX
OS Arabidopsis thaliana.
XX
PN WO9729123-A2.
XX
PD 14-AUG-1997.
XX
PF 12-FEB-1997; 97WO-GB00390.
XX
PR 12-FEB-1996; 96GB-0002796.
XX
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
XX
PI Carol P, Harberd NP, Peng J, Richards DE;
XX
WPI; 1997-415295/38.
DR N-PSDB; AAT91937.
XX
PT Nucleic acid encoding gibberellin inhibitor GAI and related

```

antisense sequences - used to create tall, or particularly, dwarf plants, especially crops such as maize, rice and wheat  
 Claim 1; Fig 4; 76pp; English.

The present sequence is the Arabidopsis thaliana gibberellin insensitivity (gai) gene product (GAI), the expression of which inhibits plant growth. However the inhibition is antagonised by gibberellin (GA), while gai expression confers a dwarf phenotype that is insensitive to GA. Manipulating gai and GAI expression can produce tall or dwarf plants, particularly the latter for increased resistance to lodging and increased yield. It may also allow regulation of flowering, i.e. plants remain in the vegetative state until treated with GA, useful to inhibit bolting in spinach and lettuce. GAI can be used to raise specific antibodies for identifying homologous proteins or genes in other species. gai fragments can also be used as probes or primers to identify and clone related sequences, or in the preparation of antisense or sense expression regulating (co-suppressing) sequences. Rice plants that express GAI may be resistant to Bakane disease. Manipulation of gai and GAI makes it possible to tailor the degree of dwarfism and GA sensitivity to particular crops or situations.

SQ Sequence 532 AA;

Query Match 86.4%; Score 70; DB 18; Length 532;

Best Local Similarity 82.4%; Pred. No. 9.7e-05; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DELLAALGKYKVRASDMA 17

Db 27 dellavlgkyvrssma 43

RESULT 11

AAE02545  
 ID AAE02545 standard; Protein: 532 AA.

XX AAC02545;

DT 10-AUG-2001 (first entry)

DE A. thaliana transcription factor G308.

KW plant transcription factor; phenotype: sugar sensing characteristic; transgenic plant; plant yield; growth; germination; photosynthesis; glyoxylate metabolism; respiration; pathogen response; wounding response; cell cycle regulation; pigmentation; flowering; senescence; physiology; storage organ; metabolism.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers  
 FT Domain 270..274  
 /note= "Conserved domain"

PN WO200135725-A1.

XX 25-MAY-2001.

PF 14-NOV-2000; 2000WO-US31414.

PR 17-NOV-1999; 99US-0166228.

PR 17-APR-2000; 2000US-0187899.

PR 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (JIAN/) JIANG C.

PA (HEAR/) HEARD J.

PA (PINE/) PINEDA O.

PA (PILG/) PILGRIM M.

PA (ADAM/) ADAM L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (SAMA/) SAMAHA R.

XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;

PI Yu G, Samaha R;

XX WPI; 2001-335977/35.

DR N-PSDB; AAD06646.

XX Nucleic acids encoding plant transcription factor polypeptides, useful for altering the sugar sensing characteristics of plants and increasing yield, e.g. corn, potato and cotton plants -

PS Claim 4; Page 74-76; 151pp; English.

XX The patent relates to polynucleotides encoding 35 plant transcription factors which may be used to modify phenotype associated with a plant's sugar sensing characteristics and increasing yield when their expression level is altered. Sugars are central regulatory molecules that control aspects of physiology, metabolism and development. Therefore the cDNAs and proteins of the invention are useful for modifying the growth and germination rates of plants, photosynthesis, glyoxylate metabolism, respiration, starch and sucrose synthesis and degradation, pathogen response, wounding response, cell cycle regulation, pigmentation, relationships in seeds, tubers, roots, and other storage organs leading to an increase in yield. The transcription factor polynucleotides and polypeptides may be used to alter the structure and developmental characteristics of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas. The present sequence is an Arabidopsis thaliana transcription factor.

SQ Sequence 532 AA;

Query Match 86.4%; Score 70; DB 22; Length 532;

Best Local Similarity 82.4%; Pred. No. 9.7e-05;

Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DELLAALGKYKVRASDMA 17

Db 27 dellavlgkyvrssma 43

RESULT 12

AAE01907

ID AAE01907 standard; Protein: 532 AA.

XX AAE01907;

DT 31-JUL-2001 (first entry)

DE Arabidopsis thaliana transcription factor, G308.

XX Transcription factor; biochemical characteristic; controlling element; structural characteristic; developmental characteristic; gene therapy; agricultural biotechnology; plant trait modification.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers  
 FT Domain 270..274  
 /label= Conserved\_domain

XX WO200136597-A1.

PN 25-MAY-2001.

PF 14-NOV-2000; 2000WO-US31344.  
 XX PR  
 PR 17-NOV-1999; 99US-0166228.  
 PR 17-APR-2000; 2000US-0197899.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX  
 (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (CREE/) CREELMAN R.  
 PA (YUGG/) YU G.  
 PA (ADAM/) ADAM L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (PILG/) PILGRIM M.  
 PA (PINE/) PINEDA O.  
 PA (JIAN/) JIANG C.  
 XX  
 Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
 PI Pilgrim M, Pineda O, Jiang C;  
 XX  
 WPI: 2001-335999/35.  
 DR N-PSDB; AAD05791.  
 XX  
 Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the biochemical characteristics of plants e.g. corn,  
 PT potato and cotton plants -  
 XX  
 PS Claim 4; Page 114-115; 127pp; English.  
 XX  
 The present sequence is Arabidopsis thaliana transcription factor,  
 CC G308, a homologue of G307. The transcription factor is used for altering  
 CC a plant's biochemical characteristics. The transcription factor may be  
 CC used to alter the structure and developmental characteristics of plants  
 CC such as soybean, wheat, corn, potato, cotton, rice, oilseed rape,  
 CC sunflower, alfalfa, sugar cane, turf, banana, blackberry, blueberry,  
 CC strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber,  
 CC eggplant, grapes, honey dew, lettuce, mango, onion, papaya, peas,  
 CC peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato,  
 CC watermelon, rosaceous fruits and/or vegetable brassicas. Transcription  
 CC factors are key controlling elements of biological pathways and altering  
 CC expression levels of 1 or more transcription factors can change entire  
 CC biological pathways in an organism. Therefore manipulating transcription  
 CC factor levels in plants offers great potential in agricultural  
 CC biotechnology for modifying a plant's traits. Transcription factor cDNA  
 CC is useful in gene therapy.  
 XX  
 SQ Sequence 532 AA;

Query Match 86.4%; Score 70; DB 22; Length 532;  
 Best Local Similarity 82.4%; Pred. No. 9.7e-05;  
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DELLALGKYVRASDMA 17  
 Db 27 dellavlgkyvrssema 43  
 RESULT 13  
 AAG38575  
 ID AAG38575 standard; Protein; 533 AA.  
 XX  
 AC AAG38575;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 47610.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.

XX EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0135124.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.



```
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.4%; Score 70; DB 21; Length 533;
Best Local Similarity 82.4%; Pred. No. 9.8e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGKVKVRSDMA 17
Db 28 dellavlgkvrsema 44

RESULT 14
AAB28574
ID AAB28574 standard; Protein; 587 AA.
XX
AC AAB28574;
XX
DT 09-FEB-2001 (first entry)
XX
DE Arabidopsis SCLa8.
XX
KW Scarecrow gene; SCR gene; plant; regulatory gene; SCAECROW-like; SCL;
transgenic plant; cell division; molecular marker; herbicide resistance;
KW salt resistance; pathogen resistance; insect resistance.
XX
OS Arabidopsis thaliana.
XX
PN WO200053723-A2.
XX
PD 14-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-US05875.
XX
```

```

PR 10-MAR-1999; 99US-0265585.
XX (UYNV ) UNIV NEW YORK STATE.
PA Benfey PN, Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysch L;
PI Helariutta Y, Bruce W, Lim J;
XX WPI; 2000-594315/56.
DR N-PSDB; AAC65291.
XX
XX Scarecrow gene useful for producing transgenic plants expressing genes
PT whose product increases starch, lignin or cellulose biosynthesis and
PT confers herbicide, pathogen or insect resistance -
XX
XX Claim 14; Fig 13; 200pp; English.
XX
XX The present sequence is given in a specification relating to the
CC structure and function of a regulatory gene, SCARECROW (SCR).
CC SCARECROW-like (SCL) genes encoding proteins containing an amino acid
CC sequence similar to the sequence of MOTIF III (WHIID) of Arabidopsis SCR
CC protein are disclosed. SCR and SCL genes are useful for producing
CC transgenic plants whose cell division is modified and root and/or stem
CC development and gravitropism of stem or hypocotyl is altered. Cell
CC division is increased or decreased in roots resulting in thicker or
CC thinner root development. The transgenic plants are useful for expressing
CC a gene of interest encoding a gene product that confers herbicide, salt,
CC pathogen or insect resistance in root or embryos and genes encoding
CC starch, lignin or cellulose biosynthesis in shoots. The SCR gene also
CC confers less susceptibility to lodging in the transgenic plants than a
CC wild-type plant. SCR gene sequences are also useful as molecular markers
CC for a quantitative trait e.g. root or gravitropism trait in molecular
XX breeding of crop plants.
XX
XX Sequence 587 AA;
SQ

```

Query Match 86.4%; Score 70; DB 21; Length 587;  
Best Local Similarity 82.4%; Pred. No. 0.00011;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 DELLAALGYKVRSDMA 17
Db 44 dellavigykvrsema 60

```

RESULT 15  
AAE02560  
ID AAE02560 standard; Protein; 587 AA.  
XX  
XX AAE02560;  
XX  
XX 10-AUG-2001 (first entry)  
XX  
XX A. thaliana transcription factor G308 homolog, G307.  
XX  
XX Plant transcription factor; phenotype; sugar sensing characteristic;  
XX transgenic plant; plant yield; growth; germination; photosynthesis;  
XX glyoxylate metabolism; respiration; pathogen response; wounding response;  
XX cell cycle regulation; pigmentation; flowering; senescence; physiology;  
XX storage organ; metabolism.  
XX  
XX Arabidopsis thaliana.  
XX  
XX Key Location/Qualifiers  
XX Domain 323..339  
XX /note= "Conserved domain"  
XX  
XX WO200135725-A1.  
XX  
XX 25-MAY-2001.  
XX  
XX 14-NOV-2000; 2000WO-US31414.  
XX

```

PR 17-NOV-1999; 99US-0166228.
PR 17-APR-2000; 2000US-0197899.
PR 22-AUG-2000; 2000US-0227439.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J.
PA (PINE/) PINEDA O.
PA (PILG/) PILGRIM M.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (SAMA/) SAMAHA R.
XX
XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;
PI Yu G, Samaha R;
XX
XX WPI; 2001-335977/35.
DR N-PSDB; AAD06661.
XX
XX Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the sugar sensing characteristics of plants and increasing
PT yield, e.g. corn, potato and cotton plants -
XX
XX Claim 4; Page 115-117; 151pp; English.
XX
XX The patent relates to polynucleotides encoding 35 plant transcription
CC factors which may be used to modify phenotype associated with a plant's
CC sugar sensing characteristics and increasing yield when their expression
CC level is altered. Sugars are central regulatory molecules that control
CC aspects of physiology, metabolism and development. Therefore the cDNAs
CC and proteins of the invention are useful for modifying the growth and
CC germination rates of plants, photosynthesis, glyoxylate metabolism,
CC respiration, starch and sucrose synthesis and degradation, pathogen
CC response, wounding response, cell cycle regulation, pigmentation,
CC flowering and senescence of plants and for modifying sink-source
CC relationships in seeds, tubers, roots, and other storage organs leading
CC to an increase in yield. The transcription factor polynucleotides and
CC polypeptides may be used to alter the structure and developmental
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.
CC The present sequence is a homolog of Arabidopsis thaliana transcription
XX factor.
XX
XX Sequence 587 AA;
SQ

```

Query Match 86.4%; Score 70; DB 22; Length 587;  
Best Local Similarity 82.4%; Pred. No. 0.00011;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 DELLAALGYKVRSDMA 17
Db 44 dellavigykvrsema 60

```

Search completed: December 19, 2001, 17:07:10  
Job time: 206 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2001, 17:05:49 ; Search time 12.86 Seconds  
(without alignments)  
100.697 Million cell updates/sec

Title: US-09-485-529-104

Perfect score: 81

Sequence: 1 DELLAALGYKVRASDMA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PTR\_68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	86.4	511	2 G96688	hypothetical prote
2	70	86.4	523	2 T51475	RGA-like protein -
3	70	86.4	533	2 H86282	protein F1086.34 [
4	70	86.4	587	2 D84426	hypothetical prote
5	46	56.8	495	2 C71679	UDP-N-acetylmurama
6	45	55.6	257	2 T10758	mandelonitrile lya
7	44	54.3	466	2 B69610	carboxy-terminal p
8	43	53.1	640	2 E64116	exodeoxyribonuclea
9	42	51.9	263	2 T01149	probable acetone-c
10	42	51.9	516	2 E96839	hypothetical prote
11	41	50.6	272	2 C72498	probable stress pr
12	41	50.6	311	2 A75047	2-ketovaleate oxi
13	41	50.6	314	2 B71114	probable ferredoxi
14	41	50.6	454	2 B70418	UDP-N-acetylmurama
15	41	50.6	463	2 D81960	ATP-dependent DNA
16	41	50.6	469	2 B81201	UDP-N-acetylmurama
17	41	50.6	484	2 C84955	UDP-N-acetylmurama
18	41	50.6	491	1 CE6CAM	UDP-N-acetylmurama
19	41	50.6	491	2 G85491	hypothetical prote
20	41	50.6	506	2 A81777	UDP-N-acetylmurama
21	41	50.6	550	2 T37519	probable amino aci
22	41	50.6	1002	2 T07292	FUN12 protein - ye
23	40	49.4	96	2 G83600	hypothetical prote
24	40	49.4	330	2 A83417	probable oxidoredu
25	40	49.4	477	2 E82763	UDP-N-acetylmurama
26	40	49.4	480	2 D83094	UDP-N-acetylmurama
27	40	49.4	700	2 T49445	adrenoleukodystrop
28	40	49.4	957	2 T10633	hypothetical prote
29	39	48.1	82	2 A72274	hypothetical prote

conserved hypothet  
probable (S)-aceto  
probable ribosomal  
hypothetical prote  
hypothetical prote  
probable acetone-c  
2,2',3-trihydroxyb  
probable molybdenu  
pyruvate synthase  
spermidine/putresc  
spermidine/putresc  
hypothetical prote  
spermidine/putresc  
probable UDP-N-ace  
hypothetical prote  
cholecystokin in ty

30 39 48.1 136 2 H75570  
31 39 48.1 141 2 T02428  
32 39 48.1 188 2 A71186  
33 39 48.1 194 2 S76919  
34 39 48.1 250 2 F83609  
35 39 48.1 263 2 T01151  
36 39 48.1 294 2 A49932  
37 39 48.1 294 2 H81448  
38 39 48.1 311 2 T45086  
39 39 48.1 377 2 F82201  
40 39 48.1 378 2 A40840  
41 39 48.1 378 2 F85683  
42 39 48.1 381 2 B64118  
43 39 48.1 604 2 C71338  
44 39 48.1 604 2 B84221  
45 39 48.1 643 2 T19135

#### ALIGNMENTS

##### RESULT 1

G96688

hypothetical protein T27F4.10 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: G96688

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salizberg, S.I.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:210161719

A:Accession: G96688

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-511 <STO>

A:Cross-references: GB:AE005173; NID:gl0092507; PIDN:AAG12907.1; GSPDB:GN00141

A:Gene: T27F4.10

A:Map position: 1

Query Match 86.4%; Score 70; DB 2; Length 511;  
Best Local Similarity 82.4%; Pred. No. 0.00017;  
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17

||||| |||||||

Db 32 DELLVWLGYKVRSSDMA 48

##### RESULT 2

T51475

RGA-like protein - Arabidopsis thaliana

N:Alternate names: protein K3M16\_60

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C:Accession: T51475

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;

submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51475

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-523 <SAT>

A:Cross-references: EMBL:AL391150

A:Experimental source: cultivar Columbia; BAC clone K3M16

C:Genetics:  
A:Map position: 5  
A:Note: K3M16\_60

Query Match 86.4%; Score 70; DB 2; Length 523;

Best Local Similarity 82.4%; Pred. No. 0.00017;  
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17

Db 34 DEFLAVLGKVRSDMA 50

RESULT 3

H86282 protein F10B6.34 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: H86282

R:Chenologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H86282

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-533 <STO>

A:Cross-references: GB:AE005172; NID:98778219; PIDN:AAF79228.1; GSPDB:GN00141

C:Genetics:

A:Gene: F10B6.34

A:Map position: 1

Query Match 86.4%; Score 70; DB 2; Length 533;

Best Local Similarity 82.4%; Pred. No. 0.00018;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17

Db 28 DEFLAVLGKVRSDMA 44

RESULT 4

DB4426

hypothetical protein At2g01570 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: DB4426

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.V.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

uss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: DB4426

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-587 <STO>

A:Cross-references: GB:AE002093; NID:93785986; PIDN:AAC67333.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g01570

A:Map position: 2

Query Match 86.4%; Score 70; DB 2; Length 587;

Best Local Similarity 82.4%; Pred. No. 0.00019;

Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17

Db 44 DEFLAVLGKVRSDMA 60

RESULT 5

C71679

UDP-n-acetylmuramate--alanine ligase (murC) RP247 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

C:Accession: C71679

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: C71679

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-495 <AND>

A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14709.1; PID:9386

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: murC; RP247

C:Superfamily: UDP-N-acetylmuramate--alanine ligase

Query Match 56.8%; Score 46; DB 2; Length 495;

Best Local Similarity 60.0%; Pred. No. 3.1;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELLAALGYKVRASDM 16

Db 31 EILHLGYKVGSD 45

RESULT 6

T10758

mandelonitrile lyase (EC 4.1.2.10) - Para rubber tree

N:Alternate names: hydroxynitrile lyase

C:Species: Hevea brasiliensis (Para rubber tree)

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: T10758

R:Hasslach, M.; Schall, M.; Hayn, M.; Griengl, H.; Kohlwein, S.D.; Schwab, H.

J. Biol. Chem. 271, 5884-5891, 1996

A:Title: Molecular-cloning of the full-length cDNA of (s)-hydroxynitrile lyase from h

n active-site residue.

A:Reference number: 203416; MUID:96215058

A:Accession: T10758

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-257 <HAS>

A:Cross-references: EMBL:U04042; NID:91223883; PIDN:AAC49184.1; PID:g1223884

A:Experimental source: leaves

C:Genetics:

C:Function:

A:Description: catalyzes the conversion of mandelonitrile to benzaldehyde and cyanide

A:Note: involved in the biodegradation of cyanogenic glycosides; also catalyzes the s

C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 55.6%; Score 45; DB 2; Length 257;

Best Local Similarity 66.7%; Pred. No. 2.4;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LLAALGYKVRASDMA 17

Db 25 LLEALGHKVTALDLA 39

RESULT 7  
B69610  
carboxy-terminal processing proteinase ctpA (EC 3.4.99.-) - Bacillus subtilis  
N:Alternate names: tail-specific endopeptidase Prc  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: B69610; JC5744  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:9804033  
A:Accession: B69610  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-466 <KUN>  
A:Cross-references: GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13850.1; PID:g2634351  
A:Experimental source: strain 168  
R:Marasco, R.; Varcamonti, M.; Ricca, E.; Sacco, M.  
Gene 183, 149-152, 1996  
A:Title: A new Bacillus subtilis gene with homology to Escherichia coli prc.  
A:Reference number: JC5744; MUID:97149292  
A:Accession: JC5744  
A:Molecule type: DNA  
A:Residues: 1-297, 'RSRNTGRC', 306-466 <MAR>  
A:Cross-references: EMBL:X98341; NID:g1402943; PIDN:CAA66987.1; PID:g1402944  
A:Experimental source: strain PY17  
C:Comment: This protein is presumed to be involved in cleavage of the carboxyl-terminal  
C:Genetics:  
A:Gene: ctpA  
A:Start codon: TTG  
C:Superfamily: carboxyl-terminal processing proteinase  
C:Keywords: hydrolase

Query Match 54.3%; Score 44; DB 2; Length 466;  
Best Local Similarity 53.3%; Pred. No. 6.6;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 2 ELAALGYKVRASDM 16  
Db 394 KMLKALGYKVKVNSM 408  
::: |||||:::  
Db 394 KMLKALGYKVKVNSM 408

RESULT 8  
E64116  
exodeoxyribonuclease V 67K chain homolog - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: E64116  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A  
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.I.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: E64116  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-640 <TTGR>  
A:Cross-references: GB:U32811; GB:L42023; NID:g1574774; PIDN:AAC22967.1; PID:g1574782; T

C:Superfamily: exodeoxyribonuclease V 67K chain

Query Match 53.1%; Score 43; DB 2; Length 640;  
Best Local Similarity 69.2%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 LLAALGYKVRASD 15  
Db 355 LLATGYKVEGSD 367  
::: |||||:::  
Db 355 LLATGYKVEGSD 367

RESULT 9  
T01149  
probable acetone-cyanohydrin lyase [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein F2686.25  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 16-Feb-2001  
C:Accession: T01149; G84626  
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke  
submitted to the EMBL Data Library, June 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F2686 genomic sequence.  
A:Reference number: Z14198  
A:Accession: T01149  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-263 <ROU>  
A:Cross-references: EMBL:AC003040; NID:g3242700; PID:g3242721  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: G84626  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-263 <SPO>  
A:Cross-references: GB:AE002093; NID:g3242721; PIDN:AAC23773.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: F2686.25; At2g23600  
A:Map position: 2  
A:Introns: 125/3; 170/3  
C:Superfamily: tropinesterase

Query Match 51.9%; Score 42; DB 2; Length 263;  
Best Local Similarity 60.0%; Pred. No. 8.4;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LLAALGYKVRASDMA 17  
Db 29 LLEALGHRVTALDLA 43  
::: |||||:::  
Db 29 LLEALGHRVTALDLA 43

RESULT 10  
E96839  
hypothetical protein F23A5.6 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: E96839  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dunaway, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizlar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719

A:Accession: E96839  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-516 <STO>  
 A:Cross-references: GB:AE005173; NID:g503282; PIDN:AAF14658.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F23A5.6  
 A:Map position: 1

Query Match 51.9%; Score 42; DB 2; Length 516;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDM 16  
 |||:||||:|:|:  
 DB 17 DEMLAALNVRKASSL 32

## RESULT 11

C72498  
 probable stress protein APE1961 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: C72498  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
 A:Reference number: A72450; MUID:99310339  
 A:Accession: C72498  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-272 <KAW>  
 A:Cross-references: DBJ:AP000063; NID:g5105654; PIDN:BAA80971.1; PID:g5105659  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE1961  
 C:Superfamily: aldehyde reductase

Query Match 50.6%; Score 41; DB 2; Length 272;  
 Best Local Similarity 37.5%; Pred. No. 13;  
 Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDM 16  
 |||:||||:|:|:  
 DB 250 DEILGSLGWRLKPEDI 265

## RESULT 12

A75047  
 2-ketoglutarate oxidoreductase chain vorb (EC 1.-.-.-) (vor) PAB1473 - Pyrococcus abyssi

C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: A75047  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
 A:Reference number: A75001  
 A:Accession: A75047  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-311 <KAW>  
 A:Cross-references: GB:AJ748287; GB:AL096836; NID:g5458657; PIDN:CAB50270.1; PID:g545878  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB1473  
 C:Superfamily: pyruvate synthase beta chain  
 C:Keywords: oxidoreductase

Query Match 50.6%; Score 41; DB 2; Length 311;

Best Local Similarity 66.7%; Pred. No. 15;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LAALGYKVRASD 15  
 ||||||:|  
 DB 90 LKALGYKVKGED 101

## RESULT 13

B71114

probable ferredoxin oxidoreductase beta subunit - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: B71114  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: B71114  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-314 <KAW>  
 A:Cross-references: NID:g3236130; PIDN:BAA29772.1; PID:g3237089  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBa  
 C:Genetics:  
 A:Gene: PH0681  
 C:Superfamily: pyruvate synthase beta chain

Query Match 50.6%; Score 41; DB 2; Length 314;  
 Best Local Similarity 66.7%; Pred. No. 15;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LAALGYKVRASD 15  
 ||||||:|  
 DB 93 LKALGYKVKGED 104

## RESULT 14

B70418

UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8) - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 16-Jul-1999  
 C:Accession: B70418  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.

Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666  
 A:Accession: B70418

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-454 <AQF>  
 A:Cross-references: GB:AE000736; NID:g2983763; PIDN:AAC07323.1; PID:g2983764; GB:AE00  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: murC  
 C:Superfamily: UDP-N-acetylmuramate--alanine ligase  
 C:Keywords: ligase

Query Match 50.6%; Score 41; DB 2; Length 454;  
 Best Local Similarity 43.8%; Pred. No. 22;  
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELAALGYKVRASDMA 17  
 ::|:|||||:|:  
 DB 24 QILLEGYKVGSDIS 39

## RESULT 15

D81960  
ATP-dependent DNA helicase NMA0433 [imported] - Neisseria meningitidis (strain Z2491 serotype 4) [NCBI]  
C:Species: Neisseria meningitidis  
C:date: 03-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: D81960  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, P.; Parkhill, J.; Jorgensen, F.L.; Haft, D.H.; White, O.; Holtzoy, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, A.L.; Rupp, B.; Salzberg, S.L.; Smith, T.F.; Staden, R.; Whitehead, I.P.  
Nature 404, 502-506, 2000  
A:title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: AB1775; GUID:20222556  
A:Accession: D81960  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-463 <PAR>  
A:cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83732.1; PID:g737918  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: hrpA'; NMA0433

```
Query Match      50.6%; Score 41; DB 2; Length 463;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

QY 2 ELLAALGYKVRASD 15  
| : : | : | | | : |  
Db 131 EIGSAVGKVRFTD 144

Search completed: December 19, 2001, 17:07:48  
Job time: 119 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2001, 17:07:34 ; Search time 10.2 seconds  
(without alignments)  
61.108 Million cell updates/sec

Title: US-09-485-529-104

Perfect score: 81

Sequence: 1 DELLAALGYKVRASDMA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	46	56.8	495	1 MURC_RICPR	Q92DS8 rickettsia
2	45	55.6	257	1 HNL_HEYBR	P52704 hevea bras
3	43	53.1	640	1 EX5A_HAEIN	P45158 haemophilus
4	41	50.6	311	1 VORE_PYRAB	Q9UVZ2 pyrococcus
5	41	50.6	311	1 VORE_PYRHO	O58414 pyrococcus
6	41	50.6	454	1 MURC_AQUAE	O67373 aquifex aeo
7	41	50.6	484	1 MURC_BUCAI	P57310 buchnera ap
8	41	50.6	491	1 MURC_ECOLI	P17952 escherichia
9	41	50.6	550	1 YAO8_SCHPO	Q10087 schizosacch
10	41	50.6	1002	1 IF2P_YEAST	P39730 saccharomyc
11	39	48.1	188	1 RL5_PYRHO	O59431 pyrococcus
12	39	48.1	293	1 DBFB_PSEPA	P47243 pseudomonas
13	39	48.1	311	1 VORE_PYRPU	O51802 pyrococcus
14	39	48.1	378	1 POT4_ECOLI	P23858 escherichia
15	39	48.1	381	1 POT4_HAEIN	P45171 haemophilus
16	39	48.1	481	1 MURC_TREPA	O83361 treponema p
17	38	46.9	273	1 FOLD_MYCGE	P47259 mycoplasma
18	38	46.9	325	1 THIL_ECOLI	P77785 escherichia
19	38	46.9	483	1 MURC_BUCAP	O51926 buchnera ap
20	38	46.9	516	1 YJJI_ECOLI	P37342 escherichia
21	38	46.9	617	1 VG31_BPMD2	O64225 mycobacteri
22	38	46.9	638	1 TKT2_BACSU	P18256 bacillus su
23	38	46.9	665	1 TKT_BUCAI	P57195 buchnera ap
24	38	46.9	705	1 RNR_AQUAE	O67834 aquifex aeo
25	38	46.9	1022	1 SC44_RICCN	Q2658 rickettsia
26	38	46.9	1079	1 IF2P_SCHPO	Q10251 schizosacch
27	38	46.9	1538	1 GLSF_PORPU	P51375 porphyra pu
28	38	46.9	3859	1 RPOA_LELV	Q04561 lelystad vi
29	37	45.7	130	1 Y73A_ARCFU	P58014 archaeoglob
30	37	45.7	146	1 YPV9_METTF	P29586 methanobact
31	37	45.7	394	1 SYVC_YEAST	P36421 saccharomyc
32	37	45.7	455	1 AROA_AGRSP	Q9r4e4 agrobacteri
33	37	45.7	515	1 UDPE_NPVSL	Q88168 spodoptera

34	37	45.7	544	1 PYRG_AZOB	P28595 azospirillum
35	37	45.7	554	1 HYES_HUMAN	P34913 homo sapien
36	37	45.7	583	1 IF2P_HALHA	O93625 halobacteri
37	37	45.7	1300	1 HRPA_ECOLI	P43329 escherichia
38	37	45.7	1304	1 HRPA_HAEIN	P45018 haemophilus
39	36	44.4	238	1 VG35_HSVII	Q00109 ictalurid h
40	36	44.4	283	1 NAT_MYCTU	P96848 mycobacteri
41	36	44.4	292	1 HEM3_METJA	O57989 methanococc
42	36	44.4	395	1 METC_BORAV	O07703 bordetella
43	36	44.4	401	1 SYVC_SCHPO	O14055 schizosacch
44	36	44.4	475	1 MURC_HAEIN	P45066 haemophilus
45	36	44.4	534	1 IF2P_SULAC	P95691 sulfolobus

#### ALIGNMENTS

RESULT	1				
MURC_RICPR	ID	MURC_RICPR	STANDARD;	PRT;	495 AA.
AC	Q92DS8;				
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	UDP-N-ACETYLURAMATE--ALANINE LIGASE (EC 6.3.2.8) (UDP-N-				
DE	ACETYLURAMOYL-L-ALANINE SYNTHETASE).				
GN	MURC OR RP247.				
OS	Rickettsia prowazekii.				
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;				
OC	Rickettsiaceae; Rickettsiae; Rickettsia.				
OX	NCBI_TaxID=782;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MADRID E;				
RX	MEDLINE=99039499; PubMed=9823893;				
RA	Andersson S.G.E., Zomorodipour A., Andersson J.O.,				
RA	Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,				
RA	Eriksson A.-S., Winkler H.H., Kurland C.G.;				
RT	"The genome sequence of Rickettsia prowazekii and the origin of				
RT	mitochondria."				
RL	Nature 396:133-140(1998).				
CC	-I- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).				
CC	-I- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLURAMOYL + L-ALANINE -				
CC	ADP + ORTHOPHOSPHATE + UDP-N-ACETYLURAMOYL-L-ALANINE.				
CC	-I- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.				
CC	-I- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).				
CC	-I- SIMILARITY: BELONGS TO THE MURDEF FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: AJ235271; CAA14709.1; -				
DR	InterPro: IPR000713; Mur_Ligase.				
DR	Pfam: PFO1225; Mur_Ligase; 1.				
KW	Peptidoglycan synthesis; Cell wall; Cell division; Ligase;				
KW	ATP-binding; Complete proteome.				
FT	NP_BIND 120 126 ATP (POTENTIAL).				
SQ	SEQUENCE 495 AA; 54612 MW; 2E18464088FAD2D6 CRC64;				

Query Match	56.8%;	Score 46;	DB 1;	Length 495;
Best Local Similarity	60.0%;	Pred. No. 1.4;		
Matches	9;	Conservative	3;	Mismatches
			3;	Indels
			0;	Gaps
			0;	
QY	2	ELLAALGYKVRASDM	16	
	1	1	1	1
Db	31	EILHNLGYKVGSDLI	45	

```
RESULT 2
HNL_HEVR
ID HNL_HEVR STANDARD; PRT; 257 AA.
AC P52704;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE (S)-ACETONE-CYANOHYDRIN LYASE (EC 4.1.2.39) ((S)-HYDROXYNITRILE LYASE)
DE ((S)-HYDROXYNITRILE) (OXYNITRILASE).
GN HNL.
OS Hevea brasiliensis (Para rubber tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eumetazoa I; Malpighiales; Euphorbiaceae; Hevea.
OX NCBI_TaxID=3981;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Leaf;
RX MEDLINE=96215058; PubMed=8621461;
RA Hasselbacher M., Schall M., Hayn M., Griengl H., Kohlwein S.D.,
RA Schwab H.;
RT "Molecular cloning of the full-length cDNA of (S)-hydroxynitrile
RT lyase from Hevea brasiliensis. Functional expression in Escherichia
RT coli and Saccharomyces cerevisiae and identification of an active
RT site residue."
RL J. Biol. Chem. 271:5884-5891(1996).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96434327; PubMed=8805565;
RA Wagner U.G., Hasselbacher M., Griengl H., Schwab H., Kratky C.;
RT "Mechanism of cyanogenesis: the crystal structure of hydroxynitrile
RT lyase from Hevea brasiliensis."
RL Structure 4:811-822(1996).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
RX MEDLINE=99423043; PubMed=10494852;
RA Gruber K., Gugganig M., Wagner U.G., Kratky C.;
RT "Atomic resolution crystal structure of hydroxynitrile lyase from
RT Hevea brasiliensis."
RL Biol. Chem. 380:993-1000(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.72 ANGSTROMS).
RX MEDLINE=20014021; PubMed=10548044;
RA Zuegg J., Gruber K., Gugganig M., Wagner U.G., Kratky C.;
RT "Three-dimensional structures of enzyme-substrate complexes of the
RT hydroxynitrile lyase from Hevea brasiliensis."
RL Protein Sci. 8:1990-2000(1999).
CC -!- FUNCTION: INVOLVED IN CYANOGENESIS, THE RELEASE OF HCN FROM
CC INJURED TISSUES. DECOMPOSES A VARIETIES OF (R) OR (S) CYNOHYDRINS
CC INTO HCN AND THE CORRESPONDING ALDEHYDES AND KETONES. THE NATURAL
CC SUBSTRATE OF THIS ENZYME IS (S)-ACETONE CYANOHYDRIN.
CC CATALYTIC ACTIVITY: (S)-2-HYDROXYISOBUTYRONITRILE - CYANIDE +
CC ACETONE.
CC -!- SUBUNIT: HOMODIMER.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: DISTANT RELATIONSHIP WITH TYPE-B
CC CARBOXYLSTERASE/LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40402; AAC49184.1;
CC PDB: 1VAS; 16-JUN-97.
CC PDB: 2VAS; 13-OCT-99.
CC PDB: 3YAS; 13-OCT-99.
CC PDB: 4YAS; 13-OCT-99.
CC PDB: 5YAS; 13-OCT-99.
CC -----
PDB; 6YAS; 13-OCT-99.
DR PDB; 1QJ4; 10-OCT-99.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000379; Est_lip_thioest_actsite.
DR Pfam: PF00561; abhydrolase; 1.
KW Lyase; 3D-structure..
FT ACT_SITE 80 80
FT ACT_SITE 207 207
FT ACT_SITE 235 235
FT MUTAGEN 80 80
FT MUTAGEN 81 81
FT MUTAGEN 235 235
SQ SEQUENCE 257 AA; 29227 MW; EF4AE88717279CEB CRC64;
Query Match 55.6%; Score 45; DB 1; Length 257;
Best Local Similarity 66.7%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 3 LLAALGYKVRASDMA 17
DB 25 LLEALGHKVTALDLA 39
RESULT 3
EX5A_HAEIN
ID EX5A_HAEIN STANDARD; PRT; 640 AA.
AC P45158;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EXODEOXYRIBONUCLEASE V ALPHA CHAIN (EC 3.1.11.5).
GN RECD OR H11322.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -!- FUNCTION: EXHIBITS SEVERAL CATALYTIC ACTIVITIES, INCLUDING
CC ATP-DEPENDENT EXONUCLEASE, ATP-STIMULATED ENDONUCLEASE,
CC ATP-DEPENDENT UNWINDING AND DNA-DEPENDENT ATPASE ACTIVITIES.
CC STRAND CLEAVAGE OCCURS 5' TO 3' DURING THE UNWINDING OF DUPLEX
CC DNA AT CHI SEQUENCES, WHICH LOCALLY STIMULATE RECOMBINATION
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: EXONUCLEOLYTIC CLEAVAGE (IN THE PRESENCE OF
CC ATP) IN EITHER 5'-TO 3'-OR 3'-TO 5'-DIRECTION TO YIELD 5'-
CC PHOSPHOOLIGONUCLEOTIDES.
CC -!- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
CC (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40402; AAC49184.1;
CC PDB: 1VAS; 16-JUN-97.
CC PDB: 2VAS; 13-OCT-99.
CC PDB: 3YAS; 13-OCT-99.
CC PDB: 4YAS; 13-OCT-99.
CC PDB: 5YAS; 13-OCT-99.
```

```
DR EMBL: U32811; AAC22967.1; -.
DR HSP: P09980; 1UAA.
DR TIGR: H11322; -.
KW Hydrolase; Nuclease; Exonuclease; Helicase; ATP-binding;
FT DNA repair; Complete proteome.
FT NP_BIND 194 201 ATP (POTENTIAL).
SQ SEQUENCE 640 AA; 72864 MW; C3C530AC398B5DA0 CRC64;

Query Match 53.1%; Score 43; DB 1; Length 640;
Best Local Similarity 69.2%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 LLAALGYKVRASD 15
   ||| |||| ||
Db 355 LLATGKVGESD 367

RESULT 4
VORB_PVRAB
ID VORB_PVRAB STANDARD; PRT; 311 AA.
AC OSU72; 2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE KETOISVALERATE OXIDOREDUCTASE SUBUNIT VORB (2-OXOISVALERATE-
DE OXOISVALERATE OXIDOREDUCTASE BETA CHAIN) (2-OXOISVALERATE-
DE FERREDOXIN OXIDOREDUCTASE BETA SUBUNIT).
GN VORB OR PAB1473.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OC NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RA Helling R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: HETEROETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE
CC GAMMA CHAIN (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ248287; CAB50270.1; -.
DR OXidoreductase; Complete proteome.
KW SEQUENCE 311 AA; 34759 MW; IEBE2609519DAC30 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 311;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LLAALGYKVRASD 15
   | ||||| |
Db 90 LKALGYKVGKED 101

RESULT 5
VORB_PVRHO
ID VORB_PVRHO STANDARD; PRT; 311 AA.
AC OS8414;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE KETOISVALERATE OXIDOREDUCTASE SUBUNIT VORB (EC 1.-.-.-) (VOR) (2-
DE OXOISVALERATE OXIDOREDUCTASE BETA CHAIN) (2-OXOISVALERATE-
```

```
DE FERREDOXIN OXIDOREDUCTASE BETA SUBUNIT).
GN VORB OR PH0681.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OC NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- SUBUNIT: HETEROETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE
CC GAMMA CHAIN (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000003; BAA29772.1; ALT_INIT.
DR OXidoreductase; Complete proteome.
KW SEQUENCE 311 AA; 34569 MW; E27B65877EF65813 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 311;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LLAALGYKVRASD 15
   | ||||| |
Db 90 LKALGYKVGKED 101

RESULT 6
MURC_AQUAE
ID MURC_AQUAE STANDARD; PRT; 454 AA.
AC O67373;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE UDP-N-ACETYLMURAMATE--ALANINE LIGASE (EC 6.3.2.8) (UDP-N-
DE ACETYLMURAMOYL-L-ALANINE SYNTHETASE).
GN MURC OR AQ1360.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OC NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLMURAMOYL + L-ALANINE =
CC ADP + ORTHOPHOSPHATE + UDP-N-ACETYLMURAMOYL-L-ALANINE.
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE000736; AAC07323.1; -;  
 DR InterPro; IPR000713; Mur\_Ligase.  
 DR Pfam; PF01225; Mur\_Ligase; 1  
 KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;  
 FT NP\_BIND 113 119 ATP (POTENTIAL).  
 SQ SEQUENCE 454 AA; 50893 MW; 12665DE14448232C CRC64;

Query Match 50.6%; Score 41; DB 1; Length 454;  
 Best Local Similarity 43.8%; Pred. No. 9.9;  
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 ELAALGYKVRASDMA 17  
 DB 24 QILLEGKVGSGSDIS 39

RESULT 7  
 MURC\_BUCAI STANDARD; PRT; 484 AA.  
 AC P57310;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE UDP-N-ACETYLURAMATE--ALANINE LIGASE (EC 6.3.2.8) (UDP-N-  
 DE ACETYLURAMOYL-L-ALANINE SYNTHETASE).  
 GN MURC OR BU215.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS";  
 RL Nature 407:81-86(2000).  
 CC -|- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).  
 CC -|- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLURAMOYL + L-ALANINE -  
 CC ADP + ORTHOPHOSPHATE + UDP-N-ACETYLURAMOYL-L-ALANINE.  
 CC -|- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS  
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
 CC -|- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AP001118; BAB12931.1; -;  
 DR InterPro; IPR001064; Crystallin.  
 DR Pfam; PF01225; Mur\_Ligase; 1  
 KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;  
 FT NP\_BIND 125 131 ATP (POTENTIAL).  
 SQ SEQUENCE 484 AA; 54311 MW; 06359D435D5590A7 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 484;  
 Best Local Similarity 57.1%; Pred. No. 11;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 LLAALGYKVRASDM 16  
 DB 37 ILLKLGKVGSGSDL 50

RESULT 8  
 MURC\_ECOLI STANDARD; PRT; 491 AA.  
 AC P17952; O07099;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE UDP-N-ACETYLURAMATE--ALANINE LIGASE (EC 6.3.2.8) (UDP-N-  
 DE ACETYLURAMOYL-L-ALANINE SYNTHETASE).  
 GN MURC OR B0091.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=90326550; PubMed=2197603;  
 RA Ikeda M., Wachi M., Jung H.K., Ishino F., Matsubashi M.;  
 RT "Nucleotide sequence involving murG and murC in the mra gene cluster  
 RT region of Escherichia coli";  
 RL Nucleic Acids Res. 18:4014-4014(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=92334977; PubMed=1630901;  
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 0-2.4 min region";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A. (MUTANT MURC3).  
 RC STRAIN-CGSC 5988;  
 RX MEDLINE=97309380; PubMed=9166795;  
 RA Eveland S.S., Pompliano D.L., Anderson M.S.;  
 RT "Conditionally lethal Escherichia coli murein mutants contain point  
 RT defects that map to regions conserved among murein and folyl poly-  
 RT gamma-glutamate ligases: identification of a ligase superfamily";  
 RL Biochemistry 36:6223-6229(1997).  
 RN [5]  
 RP SEQUENCE OF 1-14, AND CHARACTERIZATION.  
 RX MEDLINE=95324553; PubMed=7601127;  
 RA Liger D., Masson A., Blanot D., van Heijenoort J., Parquet C.;  
 RT "Over-production, purification and properties of the uridine-  
 RT diphosphate-N-acetylmuramate:L-alanine ligase from Escherichia coli";  
 RL Eur. J. Biochem. 230:80-87(1995).  
 CC -|- FUNCTION: CELL WALL FORMATION.  
 CC -|- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLURAMOYL-L-ALANINE -  
 CC ADP + ORTHOPHOSPHATE + UDP-N-ACETYLURAMOYL-L-ALANINE.  
 CC -|- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.  
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
 CC -|- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X52644; CAA36868.1; -;  
 DR EMBL; X55034; CAA38868.1; -;  
 DR EMBL; D10483; BAA01356.1; -;  
 DR EMBL; AE000118; AAC73202.1; -;  
 DR EMBL; U67892; AAB60787.1; -;  
 DR PIR; JQ0545; CEECAM.  
 DR PIR; JQ4601; S40601.  
 DR EcoGene; EGI0619; murC.  
 DR InterPro; IPR000713; Mur\_ligase.  
 DR Pfam; PF01225; Mur\_ligase; 1.  
 KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;  
 KW ATP-binding; Complete proteome.  
 FT NP\_BIND 126 132 ATP (POTENTIAL).  
 FT MUTAGEN 344 344 G->D: IN MURC3.  
 SQ SEQUENCE 491 AA; 53626 MW; D201B35931C013FB CRC64;

Query Match 50.6%; Score 41; DB 1; Length 491;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELIAALGYKVRSDMA 17  
 Db 37 EVLANEGYQISGSDLA 52

RESULT 9  
 ID YAO8\_SCHPO STANDARD; PRT; 550 AA.  
 AC Q10087;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PUTATIVE AMINO-ACID PERMEASE CLID3.08C.  
 GN SPAC11D3.08C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z68166; CAA92309.1; -;  
 DR InterPro; IPR002293; AA\_rel\_permease\_1.  
 DR InterPro; IPR002027; Amino\_acid\_permease.  
 DR Pfam; PF00324; aa\_permeases; 1.  
 DR PROSITE; PS00218; AMINO-ACID\_PERMEASE\_1; 1.  
 KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane.  
 FT TRANSMEM 53 73 POTENTIAL.  
 FT TRANSMEM 78 98 POTENTIAL.  
 FT TRANSMEM 127 147 POTENTIAL.

FT TRANSMEM 177 197 POTENTIAL.  
 FT TRANSMEM 201 221 POTENTIAL.  
 FT TRANSMEM 246 266 POTENTIAL.  
 FT TRANSMEM 283 303 POTENTIAL.  
 FT TRANSMEM 336 356 POTENTIAL.  
 FT TRANSMEM 386 406 POTENTIAL.  
 FT TRANSMEM 411 431 POTENTIAL.  
 FT TRANSMEM 453 473 POTENTIAL.  
 FT TRANSMEM 481 501 POTENTIAL.  
 SQ SEQUENCE 550 AA; 59918 MW; 3DD1EF6003896EAC CRC64;

Query Match 50.6%; Score 41; DB 1; Length 550;  
 Best Local Similarity 80.0%; Pred. No. 12;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYK 10  
 Db 32 DELLMSLGYK 41

RESULT 10  
 ID IF2P\_YEAST STANDARD; PRT; 1002 AA.  
 AC P39730;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE TRANSLATION INITIATION FACTOR IF-2.  
 GN YAL035W OR FUN12.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=95249563; PubMed=7731988;  
 RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,  
 RA Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,  
 RA Storms R.K.;  
 RT "The nucleotide sequence of chromosome I from Saccharomyces  
 RT cerevisiae".  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94357438; PubMed=8076820;  
 RA Suttrave P., Shafer B.K., Strathern J.N., Hughes S.H.;  
 RT "Isolation, identification and characterization of the FUN12 gene of  
 RT Saccharomyces cerevisiae".  
 RL Gene 146:209-213(1994).  
 RN [3]  
 RP REVISIONS.  
 RC STRAIN=S288C / AB972;  
 RA Vo D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC CHARACTERIZATION.  
 RX MEDLINE=98288357; PubMed=9624054;  
 RA Choi S.K., Lee J.H., Zoll W.L., Merrick W.C., Dever T.E.;  
 RT "Promotion of met-tRNA<sup>Met</sup> binding to ribosomes by yif2, a bacterial  
 RT IF2 homolog in yeast".  
 RL Science 280:1757-1760(1998).  
 CC -1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING  
 CC THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO  
 CC FUNCTION ALONG WITH EIF-2.  
 CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U12980; AAC04996.1; -;  
 DR EMBL; L29389; AAA57228.1; ALT\_SEQ.  
 DR SGD; S0000033; FUN12.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR00178; IF2.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF02131; IF2; 1.  
 DR Initiation factor; Protein biosynthesis; GTP-binding.  
 KW DOMAIN 361 371 POLY-GLU.  
 FT NP\_BIND 412 419 GTP (BY SIMILARITY).  
 SQ SEQUENCE 1002 AA; 112268 MW; 1A496195DAE1C283 CRC64;

Query Match 50.68; Score 41; DB 1; Length 1002;  
 Best Local Similarity 53.38; Pred. No. 22;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELLAALGYKVRASDM 16  
 I: ||||| : | : | :  
 Db 698 EVKALGVKIAANDL 712

RESULT 11  
 RL5\_PYRHO STANDARD; PRT; 188 AA.  
 AC OS9431;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L5P.  
 GN RPL5P OR PH1765.  
 GE Pyrococcus horikoshii.  
 OS Pyrococcus horikoshii.  
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shiraya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AP000007; BAA30880.1; -;  
 DR InterPro; IPR002132; Ribosomal\_L5.  
 DR Pfam; PF00281; Ribosomal\_L5; 1.  
 DR Pfam; PF00673; Ribosomal\_L5\_C; 1.  
 DR ProDom; PD001076; Ribosomal\_L5; 1.  
 DR PROSITE; PS00356; RIBOSOMAL\_L5; FALSE\_NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 188 AA; 21610 MW; 56CAF0B5053AECF4 CRC64;

Query Match 48.18; Score 39; DB 1; Length 188;  
 Best Local Similarity 66.78; Pred. No. 9.2;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LLAALGYKVRAS 14  
 ||||| : | : | :  
 Db 94 LLAADVYKIKAS 105

RESULT 12  
 DBFB\_PSEPA STANDARD; PRT; 293 AA.  
 AC PA7243;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 2,2',3'-TRIHYDROXYBIPHENYL DIOXYGENASE (EC 1.13.11.-).  
 GN DBFB.  
 OS Pseudomonas paucimobilis (Sphingomonas paucimobilis).  
 CC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
 OC Sphingomonas.  
 OX NCBI\_TaxID=13689;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-63.  
 RC STRAIN=RW1;  
 RX MEDLINE=94042906; PubMed=8226678;  
 RA Happe B., Eltis L.D., Poth H., Hedderich R., Timmls K.N.;  
 RT "Characterization of 2,2',3'-trihydroxybiphenyl dioxygenase, an  
 RT extradiol dioxygenase from the dibenzofuran- and  
 RT dibenzo-p-dioxin-degrading bacterium Sphingomonas sp. strain RW1.";  
 RL J. Bacteriol. 175:7313-7320(1993).  
 RN [2]

RP REVISIONS.  
 RA Armengaud J.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: RESPONSIBLE FOR META CLEAVAGE OF THE FIRST AROMATIC RING  
 CC OF 2,2',3'-TRIHYDROXYBIPHENYL AND 2,3-DIHYDROXYBIPHENYL 2,2',3-  
 CC TRIHYDROXYDIPHENYL ETHER, CATECHOL, 3-METHYL-CATECHOL, AND 4-  
 CC METHYL-CATECHOL ARE OXIDIZED LESS EFFICIENTLY AND 3,4-  
 CC DIHYDROXYBIPHENYL IS OXIDIZED CONSIDERABLY LESS EFFICIENTLY.  
 CC -1- COFACTOR: FERROUS ION.  
 CC -1- PATHWAY: KEY ENZYME IN THE DEGRADATION PATHWAYS OF DIBENZO-P-  
 CC DIOXIN AND DIBENZOFURAN (SECOND STEP).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X72850; CAA51364.1; -;  
 DR HSP; P47228; IHAN.  
 DR InterPro; IPR000486; Extradiol\_dioxygenase.  
 DR InterPro; IPR000325; Glyoxalase\_1.  
 DR Pfam; PF00903; Glyoxalase; 2.  
 DR ProDom; PD000977; Extradiol\_dioxygenase; 1.  
 DR PROSITE; PS00082; EXTRADIOL\_DIOXYGENAS; 1.  
 KW Oxidoreductase; dioxygenase; Aromatic hydrocarbons catabolism; Iron.  
 FT INIT\_MET 0  
 FT METAL 146 146 IRON (BY SIMILARITY).  
 FT METAL 208 208 IRON (BY SIMILARITY).  
 FT METAL 259 259 IRON (BY SIMILARITY).  
 SQ SEQUENCE 293 AA; 32146 MW; E77B67329C87B32 CRC64;

Query Match 48.18; Score 39; DB 1; Length 293;  
 Best Local Similarity 57.18; Pred. No. 14;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LLAALGYKVRASDMA 17  
 ||||| : | : | :  
 Db 94 LLAADVYKIKAS 105

Db 61 LAAGNWDPSDLA 74

RESULT 13  
VORR\_PVRFU  
ID VORR\_PVRFU STANDARD; PRT; 311 AA.  
AC Q51802;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE KETOISOVALERATE OXIDOREDUCTASE SUBUNIT VORR (EC 1.-.-.-) (VOR) (2-  
DE OXISOVALERATE OXIDOREDUCTASE BETA CHAIN) (2-OXISOVALERATE-  
DE FERREDOXIN OXIDOREDUCTASE BETA SUBUNIT).  
OS VORR.  
GN Pyrococcus furiosus.  
OC Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
OX NCBI\_TaxID=2261;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-7.  
RC STRAIN=DSM 3638;  
RX MEDLINE=96125254; PubMed=8550425;  
RA Kletzin A., Adams M.W.A.;  
RT "Molecular and phylogenetic characterization of pyruvate and 2-  
RT ketoisovalerate ferredoxin oxidoreductases from Pyrococcus furiosus  
RT and pyruvate ferredoxin oxidoreductase from Thermotoga maritima";  
RL J. Bacteriol. 178:248-257(1996).  
CC -1- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE  
CC GAMMA CHAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X85250; CAA59503.1; -.  
KW Oxidoreductase.  
SQ SEQUENCE 311 AA; 34766 MW; 07CC02452E3C6074 CRC64;  
  
Query Match 48.1%; Score 39; DB 1; Length 311;  
Best Local Similarity 58.3%; Pred. No. 15;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 4 LAALGYKVRASD 15  
| : ||||| :  
Db 90 LKAMGYKVKGED 101  
  
RESULT 14  
POTA\_ECOLI  
ID POTA\_ECOLI STANDARD; PRT; 378 AA.  
AC P23858;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA.  
GN POTA OR B1126 OR Z1831 OR ECS1571.  
OS Escherichia coli, and  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562, 83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92041956; PubMed=1939142;  
RA Furuchi T., Kashiwagi K., Kobayashi H., Igarashi K.;  
RT "Characteristics of the gene for a spermidine and putrescine  
RT transport system that maps at 15 min on the Escherichia coli  
RT chromosome.";  
J. Biol. Chem. 266:20928-20933(1991).  
[2]  
SEQUENCE FROM N.A.  
RN STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
Science 277:1453-1474(1997).  
RL Science 277:1453-1474(1997).  
RN [3]  
SEQUENCE FROM N.A.  
RP STRAIN=K12;  
RX MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
RA Yano M., Horiuchi T.;  
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
RT corresponding to the 12.7-28.0 min region on the linkage map.";  
DNA Res. 3:137-155(1996).  
RN [4]  
SEQUENCE FROM N.A.  
RP STRAIN=0157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
Nature 409:529-533(2001).  
RN [5]  
SEQUENCE FROM N.A.  
RP STRAIN=0157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
DNA Res. 8:11-22(2001).  
RL DNA Res. 8:11-22(2001).  
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
CC COUPLING TO THE TRANSPORT SYSTEM.  
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED.  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M64519; AAC37038.1; -.  
DR EMBL; AE000212; AAC74210.1; -.  
DR EMBL; D90747; BAA35946.1; -.  
DR EMBL; D90748; BAA35948.1; -.  
DR EMBL; AE005326; AAG55930.1; -.  
DR EMBL; AP002555; BAB34994.1; -.  
DR PIR; A40840; A40840.  
DR Ecogene; EG10749; potA.  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR001687; ATP\_GTP\_A.



DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW Transport; ATP-binding; Inner membrane; Complete proteome.  
FT NP\_BIND 50 57 ATP (BY SIMILARITY).  
SQ SEQUENCE 378 AA; 43028 MW; 14DCA99329A344F3 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 378;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELAALGYKVR 12  
|||:|||||  
Db 172 DESLSALDYKLR 183

## RESULT 15

POTA\_HAEIN STANDARD; PRT; 381 AA.  
AC P45171;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA.  
GN POTA OR H11347  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerkavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McInnes K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
RT influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32813; AAC22991.1; -;  
DR TIGR; H11347; -;  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW Transport; ATP-binding; Inner membrane; Complete proteome.  
FT NP\_BIND 52 59 ATP (POTENTIAL).  
SQ SEQUENCE 381 AA; 43429 MW; 3E61DD3D062EBE6A CRC64;

Query Match 48.1%; Score 39; DB 1; Length 381;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELAALGYKVR 12  
|||:|||||  
Db 174 DESLSALDYKLR 185

Search completed: December 19, 2001, 17:10:29  
Job time: 175 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2001, 17:07:14 ; Search time 22.97 Seconds  
(without alignments)  
108.256 Million cell updates/sec

Title: US-09-485-529-104  
Perfect score: 81  
Sequence: 1 DELLAALGYKVRASDMA 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_17:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	81	100.0	623	10	Q9ST59
2	78	96.3	625	10	Q9MB96
3	78	96.3	630	10	Q9ST48
4	70	86.4	511	10	Q9C8Y3
5	70	86.4	523	10	Q9LF53
6	70	86.4	532	10	Q23643
7	70	86.4	532	10	Q23724
8	70	86.4	533	10	Q9LQT8
9	70	86.4	547	10	Q9SRP9
10	70	86.4	587	10	Q23725
11	70	86.4	587	10	Q23642
12	70	86.4	587	10	Q9SLH3
13	70	86.4	662	10	O65367
14	44	54.3	411	4	O43916
15	44	54.3	466	2	O45645
16	44	54.3	466	2	O34666
17	44	54.3	1149	12	Q9WBQ8
18	44	54.3	1457	12	Q9WJB3
19	44	54.3	1457	12	Q9WBQ4

20	44	54.3	1457	12	Q9DY97
21	44	54.3	1463	12	Q9YN01
22	44	54.3	1463	12	Q9J7C0
23	44	54.3	1463	12	Q9ENK5
24	44	54.3	1463	12	Q9E8M9
25	44	54.3	1463	12	Q99BU5
26	44	54.3	1463	12	Q99AV5
27	44	54.3	3956	12	Q9DLN9
28	44	54.3	3960	12	Q9DLP1
29	44	54.3	3960	12	Q9DLP0
30	44	54.3	3960	12	Q9DLN8
31	42	51.9	263	10	O80476
32	42	51.9	516	10	Q9SAI7
33	41	50.6	217	2	P70930
34	41	50.6	272	1	Q9YAH8
35	41	50.6	463	2	Q9JWD2
36	41	50.6	467	2	O30707
37	41	50.6	469	2	Q9K0Y1
38	41	50.6	506	2	Q9JSZ8
39	41	50.6	999	12	P91577
40	41	50.6	1411	12	Q06503
41	40	49.4	96	2	Q916C5
42	40	49.4	330	2	Q912R2
43	40	49.4	477	2	Q9PF80
44	40	49.4	480	2	Q9HW02
45	40	49.4	487	2	Q9GR66

ALIGNMENTS

RESULT 1

Q9ST59 PRELIMINARY; PRT; 623 AA;

AC Q9ST59; (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE GIBBERELLIN RESPONSE MODULATOR.

GN RHT-DIA.

OS Triticum aestivum (Wheat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;

OC Triticeae; Triticum.

OX NCBI\_TaxID=4565;

RP [1]

RN SEQUENCE FROM N.A.

RA MEDLINE=99347734; PubMed=10421366;

RA Peng J., Richards D.E., Hartley N.M., Murphy G.P., Devos K.M.,

RA Flintham J.E., Beales J., Fish L.J., Worland A.J., Pelica F.,

RA Sudhakar D., Christou P., Snape J.W., Gale M.D., Harberd N.P.;

RT "Green revolution" genes encode mutant gibberellin response

RL modulators.;

RL Nature 400:256-261(1999).

DR EMBL; AJ242531; CAB51555.1; .

SQ SEQUENCE 623 AA; 65337 MW; F2BAC34996D0A84F CRC64;

Query Match 100.0%; Score 81; DB 10; Length 623;  
Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
Db 38 DELLAALGYKVRASDMA 54

RESULT 2

Q9MB96 PRELIMINARY; PRT; 625 AA.

ID Q9MB96

AC Q9MB96; (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
DE OSGAI.  
GN Oryza sativa (Rice).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoideae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RX MEDLINE=20179680; PubMed=10713441;  
RA Ogawa M., Kusano T., Katsumi M., Sano H.;  
RT "Rice gibberellin-insensitive gene homolog, OSGAI, encodes a nuclear-  
RT localized protein capable of gene activation at transcriptional  
RT level";  
RL Gene 245:21-29(2000).  
DR EMBL; AB030956; BAA90749.1; -.  
SQ SEQUENCE 625 AA; 65406 MW; 034FF02719D42E97 CRC64;

Query Match 96.3%; Score 78; DB 10; Length 625;  
Best Local Similarity 94.1%; Pred. No. 7.2e-05;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17  
DB 39 DELLAALGYKVRSDMA 55  
|||||  
RESULT 3  
Q9ST48 PRELIMINARY; PRT; 630 AA.  
AC Q9ST48;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE GIBBERELLIN RESPONSE MODULATOR (FRAGMENT).  
GN D8.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99347734; PubMed=10421366;  
RA Peng J., Richards D.E., Hartley N.M., Murphy G.P., Devos K.M.,  
RA Flintham J.E., Beales J., Fish L.J., Worland A.J., Pelica F.,  
RA Sudhakar D., Christou P., Snape J.W., Gale M.D., Harberd N.P.;  
RT "Green revolution" genes encode mutant gibberellin response  
RT modulators";  
RL Nature 400:256-261(1999).  
DR EMBL; AJ242530; CAB51557.1; -.  
FT NON\_TER 630  
SQ SEQUENCE 630 AA; 66028 MW; 3D56851726C51042 CRC64;

Query Match 96.3%; Score 78; DB 10; Length 630;  
Best Local Similarity 94.1%; Pred. No. 7.3e-05;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17  
DB 38 DELLAALGYKVRSDMA 54  
|||||  
RESULT 4  
Q9CBY3 PRELIMINARY; PRT; 511 AA.  
AC Q9CBY3;  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE GIBBERELLIN REGULATORY PROTEIN, PUTATIVE.  
GN T27F4.10.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
RT thaliana";  
RL Nature 408:816-820(2000).  
DR EMBL; AC020665; AAG52171.1; -.  
SQ SEQUENCE 511 AA; 56754 MW; 1560071697C92A9F CRC64;

Query Match 86.4%; Score 70; DB 10; Length 511;  
Best Local Similarity 82.4%; Pred. No. 0.0013;  
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRSDMA 17  
DB 32 DELLVVLGYKVRSDMA 48  
|||||  
RESULT 5  
Q9LF53 PRELIMINARY; PRT; 523 AA.  
AC Q9LF53;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
DE RGA-LIKE PROTEIN.  
GN K3M16\_60.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,  
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL391150; CAC01893.1; -.  
SQ SEQUENCE 523 AA; 57326 MW; 0F6CE0BD13403C35 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 523;  
Best Local Similarity 82.4%; Pred. No. 0.0014;

Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
 ||||| |||||:||||  
 Db 34 DELAVLGYKVRSSDMA 50

RESULT 6  
 O23643 PRELIMINARY; PRT; 532 AA.  
 AC O23643;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE RGA2 PROTEIN.  
 GN RGA2  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97379310; PubMed=9237632;  
 RA Truong H.N., Caboche M., Daniel-Vedele F.;  
 RT "Sequence and characterization of two Arabidopsis thaliana CDNAS  
 RT isolated by functional complementation of a yeast gln3 gdh1 mutant.";  
 RL FEBS Lett. 410:213-218(1997).  
 DR EMBL; Y11337; CAA72178.1; -;  
 DR Mendel; 24146; Arath:3051;24146.  
 SQ SEQUENCE 532 AA; 58789 MW; 17AC719CA8072239 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 532;  
 Best Local Similarity 82.4%; Pred. No. 0.0014;  
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
 ||||| |||||:||||  
 Db 27 DELAVLGYKVRSSDMA 43

RESULT 7  
 O23724 PRELIMINARY; PRT; 532 AA.  
 AC O23724;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE GAI PROTEIN.  
 GN GAI.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LANDSBERG ERECTA;  
 RX MEDLINE=98051192; PubMed=9389651;  
 RA Peng J., Carol P., Richards D.E., King K.E., Cowling R.J.,  
 RA Murphy G.P., Harberd N.P.;  
 RT "The Arabidopsis GAI gene defines a signaling pathway that negatively  
 RT regulates gibberellin responses.";  
 RL Genes Dev. 11:3194-3205(1997).  
 DR EMBL; Y15193; CAA75492.1; -;  
 DR Mendel; 24070; Arath:3051;24070.  
 SQ SEQUENCE 532 AA; 58789 MW; 4AF4BC6EC4265503 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 532;  
 Best Local Similarity 82.4%; Pred. No. 0.0014;  
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
 ||||| |||||:||||  
 Db 27 DELAVLGYKVRSSDMA 43

RESULT 8  
 Q9LQT8 PRELIMINARY; PRT; 533 AA.  
 AC Q9LQT8;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE F10B6.34.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chao Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,  
 RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,  
 RA Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,  
 RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,  
 RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome  
 RT 1.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu S., Liu J., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006917; AAF79228.1; -;  
 SQ SEQUENCE 533 AA; 58926 MW; 6464B8C129D79528 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 533;  
 Best Local Similarity 82.4%; Pred. No. 0.0014;  
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17  
 ||||| |||||:||||  
 Db 28 DELAVLGYKVRSSDMA 44

RESULT 9  
 Q9SRP9 PRELIMINARY; PRT; 547 AA.  
 AC Q9SRP9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE RGA2-LIKE PROTEIN.  
 GN T21P5.13.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
  Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
  Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T21P5 genomic sequence.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC009895; AAF01590.1; --
SQ SEQUENCE 547 AA; 60493 MW; C4D18D5951D95634 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 547;
Best Local Similarity 82.4%; Pred. No. 0.0014;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17
Db 44 DELLAVLGYKVRSEMA 60
||||| |||||:|:|

RESULT 10
O23725 ID O23725 PRELIMINARY; PRT; 587 AA.
AC O23725;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GRS PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Peng J., Carol P., Richards D.E., King K.E., Cowling R.J.,
  Murphy G.P., Harberd N.P.;
RL Genes Dev. 0:0-0(0)
DR EMBL: Y15194; CAA75493.1; --
DR Mendel: 24071; Arath;3051;24071.
SQ SEQUENCE 587 AA; 64006 MW; F6F6C7738EE7DCA9 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 587;
Best Local Similarity 82.4%; Pred. No. 0.0015;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17
Db 44 DELLAVLGYKVRSEMA 60
||||| |||||:|:|

RESULT 11
O23642 ID O23642 PRELIMINARY; PRT; 587 AA.
AC O23642;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE RGAL PROTEIN.
GN RGAL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97379310; PubMed=9237632;
RA Truong H.N., Caboche M., Daniel-Vedele F.;

RT "Sequence and characterization of two Arabidopsis thaliana cDNAs
  isolated by functional complementation of a yeast gln3 gdh1 mutant.";
RT FEBS Lett. 410:213-218(1997).
DR EMBL: Y11336; CAA72177.1; --
DR Mendel: 24145; Arath;3051;24145.
SQ SEQUENCE 587 AA; 64023 MW; DOA7A3C741FB51EF CRC64;

Query Match 86.4%; Score 70; DB 10; Length 587;
Best Local Similarity 82.4%; Pred. No. 0.0015;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17
Db 44 DELLAVLGYKVRSEMA 60
||||| |||||:|:|

RESULT 12
O9SLH3 ID O9SLH3 PRELIMINARY; PRT; 587 AA.
AC O9SLH3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PUTATIVE RGAL, GIBERELLIN RESPONSE MODULATION PROTEIN.
GN AT2G01570.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
  Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
  Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
  Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
  Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
  Salzberg S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL: AC005560; AAC67333.1; --
SQ SEQUENCE 587 AA; 64035 MW; FC92E7F9408072AA CRC64;

Query Match 86.4%; Score 70; DB 10; Length 587;
Best Local Similarity 82.4%; Pred. No. 0.0015;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DELLAALGYKVRASDMA 17
Db 44 DELLAVLGYKVRSEMA 60
||||| |||||:|:|

RESULT 13
O65367 ID O65367 PRELIMINARY; PRT; 662 AA.
AC O65367;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RGAL PROTEIN.
GN RGA-LIKE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=CV. COL-0;
RA Sanchez-Fernandez R., Ardiiles-Diaz W., van Montagu M., Inze D.,
RA May M.J.;
RT "Cloning of the Arabidopsis thaliana RGA-like gene, a putative member
RT of the VHIID domain transcription factor family.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ224957; CAA12242.1; -.
DR Mendel; 29006; Arath;3051;29006.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 1.
DR SMART; SM00320; WD40; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 662 AA; 73126 MW; 958AB73D42121CA8 CRC64;

Query Match 86.4%; Score 70; DB 10; Length 662;
Best Local Similarity 82.4%; Pred. No. 0.0018;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DELLAALGYKVRASDMA 17
Db 183 DELLVVGKVRSSDMA 199

RESULT 14
ID O43916 PRELIMINARY; PRT; 411 AA.
AC O43916;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE CHONDROITIN-6-SULFOTRANSFERASE.
GN CHST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97189336; PubMed=9117134;
RA Williams K.J.;
RT "Atherosclerosis: cell biology and lipoproteins.";
RL Curr. Opin. Lipidol. 7:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Peng T., Tabas I., Williams K.J.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98070405; PubMed=9405439;
RA Fukuta M., Inazawa J., Torii T., Tsuzuki K., Shimada E., Habuchi O.;
RT "Molecular cloning and characterization of human keratan sulfate
RT Gal-6-sulfotransferase.";
RL J. Biol. Chem. 272:32321-32328(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL VEIN ENDOTHELIUM;
RX MEDLINE=99168906; PubMed=10049591;
RA Li X., Tedder T.F.;
RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
RT endothelial cells: cDNA cloning, expression, and chromosomal
RT localization.";
RL Genomics 55:345-347(1999).
DR EMBL; U65637; AAC28776.1; -.
DR EMBL; AB003791; BAA24840.1; -.
DR EMBL; AF090137; AAD19878.1; -.
KW Transferase.
SQ SEQUENCE 411 AA; 46714 MW; F07D0A23B6338A09 CRC64;
```

```
Query Match 54.3%; Score 44; DB 4; Length 411;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELLAALGYKVRASD 15
Db 378 QVLAQLGYKIAASE 391

RESULT 15
ID O45645 PRELIMINARY; PRT; 466 AA.
AC O45645;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ORF01 DNA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PY17;
RX MEDLINE=97149292; PubMed=8996100;
RA Marasco R., Varcamonti M., Ricca E., Sacco M.;
RT "A new Bacillus subtilis gene with homology to Escherichia coli prc.";
RL Gene 183:149-152(1996).
DR EMBL; X98341; CAA66987.1; -.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR003581; TSpc.
DR Pfam; PF00595; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00245; TSpc; 1.
SQ SEQUENCE 466 AA; 51413 MW; 4430B459C45E8714 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 466;
Best Local Similarity 53.3%; Pred. No. 32;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELLAALGYKVRASDM 16
Db 394 KMLKALGYKVKVNSM 408
```

Search completed: December 19, 2001, 17:10:13  
Job time: 179 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 17:30:48 ; Search time 1422.71 Seconds  
(without alignments)  
591.376 Million cell updates/sec

Title: US-09-485-529-105  
Perfect score: 51  
Sequence: 1 gacgagctgctggcgcgct.....tcgcgctcogacatggcg 51

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pi:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rod:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	100.0	51	6	AX005896	AX005896 Sequence
2	51	100.0	453	6	AX005857	AX005857 Sequence
3	51	100.0	1872	8	TAE242531	AJ242531 Triticum
4	51	100.0	2125	6	AX005805	AX005805 Sequence
5	51	100.0	2709	6	AX005794	AX005794 Sequence
6	49	96.1	324	6	AX005867	AX005867 Sequence
7	45.2	88.6	770	6	AX005803	AX005803 Sequence
8	45.2	88.6	2500	8	AB030956	AB030956 Oryza sat
9	45.2	88.6	122497	8	AC087797	AC087797 Oryza sat
10	44.2	86.7	725	6	AX005810	AX005810 Sequence
11	43.2	84.7	200	6	AX005860	AX005860 Sequence
12	39.8	78.0	302	6	AX005807	AX005807 Sequence
13	39.8	78.0	371	6	AX005808	AX005808 Sequence
14	39.8	78.0	1890	8	ZMA242530	AJ242530 Zea mays
15	39.8	78.0	2255	6	AX005806	AX005806 Sequence
16	39	76.5	416	6	AX005809	AX005809 Sequence
17	29.8	58.4	1542	1	AF071051	AF071051 Streptomy
18	29.8	58.4	1542	6	AR151707	AR151707 Sequence
19	29.8	58.4	1542	6	AX006515	AX006515 Sequence
20	29.8	58.4	6584	1	SCPACAS	X84101 S.clavulige
21	29.8	58.4	11604	6	AR151704	AR151704 Sequence
22	29.8	58.4	15079	6	AR151702	AR151702 Sequence
23	29.8	58.4	15120	1	SCU87786	U87786 Streptomyce
24	29.4	57.6	309	6	AX005855	AX005855 Sequence
25	28.6	56.1	230	6	AX005861	AX005861 Sequence
26	28.2	55.3	2295	1	STMTRIP	L27466 Streptomyce
27	28.2	55.3	26555	1	SC2G1	AL391014 Streptomy
28	27.8	54.5	1691	8	AB057426	AB057426 Coriolus
29	27.6	54.1	4081	8	ATA224957	AJ224957 Arabidops
30	27.6	54.1	26604	8	ATK3M16	AL391150 Arabidops
31	27.6	54.1	45335	1	SCC57A	AL136519 Streptomy
32	27.6	54.1	82289	8	ATAC009895	AC009895 Arabidops
33	27.6	54.1	85702	8	AC020665	AC020665 Arabidops
34	27.6	54.1	132699	8	AC006917	AC006917 Genomic s
35	27.4	53.7	285	6	AX005865	AX005865 Sequence
36	27	52.9	4523	14	MMULGENES	X95710 Pseudorabie
37	26.6	52.2	345783	1	AP003001	AP003001 Mesorhizo
38	26	51.0	1951	8	ATRG42	Y11337 A.thaliana
39	26	51.0	1964	6	A64697	A64697 Sequence 1
40	26	51.0	1964	8	ATY15193	Y15193 Arabidops
41	26	51.0	11093	1	AE005134	AE005134 Halobacte
42	26	51.0	11925	1	AE005111	AE005111 Halobacte
43	26	51.0	12510	1	AE005848	AE005848 Caulobact
44	26	51.0	33320	1	SC9B10	AL003204 Streptomy
45	26	51.0	37000	14	BHTIUL	Z78205 Bovine herp

ALIGNMENTS

RESULT 1

AX005896  
LOCUS AX005896 51 bp DNA  
DEFINITION Sequence 105 from Patent WO9909174.  
ACCESSION AX005896  
VERSION AX005896.1 GI:9928885  
KEYWORDS  
SOURCE  
ORGANISM

bread wheat.  
Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

1 (bases 1 to 51)

REFERENCE  
AUTHORS Harberd,N.P. and Peng,J.  
TITLE Genetic control of plant growth and development  
JOURNAL Patent: WO 9909174-A 105 25-FEB-1999;

FEATURES  
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)

Location/Qualifiers

source

1. .51  
/organism="Triticum aestivum"

```

BASE COUNT      7 a      16 c      21 g      7 t
ORIGIN

Query Match      100.0%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacagctgctggcgccgctcggtgacaaagtgccgctccgacatggcg 51
|||||
Db 1 GACGAGCTGCTGGCGCGCCTCGGTACAAAGTGCGCGCTCCGACATGGCG 51

RESULT 2
LOCUS      AX005857      453 bp      DNA      PAT      24-AUG-2000
DEFINITION      Sequence 66 from Patent WO9909174;
ACCESSION      AX005857
VERSION      AX005857.1      GI:9928852
KEYWORDS      bread wheat.
SOURCE      Triticum aestivum
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 453)
AUTHORS      Harberd,N.P. and Peng,J.
TITLE      Genetic control of plant growth and development
JOURNAL      Patent: WO 9909174-A 66 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES      source
1..453
/organism="Triticum aestivum"
/db_xref="taxon:4565"
misc_feature      6
/note="n is any nucleotide"
misc_feature      11
/note="n is any nucleotide"
misc_feature      46
/note="n is any nucleotide"
misc_feature      85
/note="n is any nucleotide"
misc_feature      107
/note="n is any nucleotide"
misc_feature      445
/note="n is any nucleotide"
misc_feature      449
/note="n is any nucleotide"
BASE COUNT      85 a      136 c      158 g      67 t
ORIGIN

Query Match      100.0%; Score 51; DB 6; Length 453;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacagctgctggcgccgctcggtgacaaagtgccgctccgacatggcg 51
|||||
Db 273 GACGAGCTGCTGGCGCGCCTCGGTACAAAGTGCGCGCTCCGACATGGCG 323

RESULT 3
LOCUS      TAE242531      1872 bp      DNA      PLN      28-JUL-1999
DEFINITION      Triticum aestivum rht-Dla gene for gibberellin response modulator.
ACCESSION      AJ242531
VERSION      AJ242531.1      GI:5640156
KEYWORDS      gibberellin response modulator; rht-Dla gene.
SOURCE      bread wheat.
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

```

Pooideae; Triticeae; Triticum.
1 (bases 1 to 1872)
Peng,J., Richards,D.E., Hartley,N.M., Murphy,G.P., Devos,K.M.,
Flintham,J.E., Beales,J., Fish,L.J., Worland,A.J., Pellica,F.,
Sudhakar,D., Christou,P., Snape,J.W., Gale,M.D. and Harberd,N.P.
'Green revolution' genes encode mutant gibberellin response
modulators
Nature 400 (6741), 256-261 (1999)
99347734
2 (bases 1 to 1872)
Harberd,N.P., Peng,J. and Richards,D.E.
Green revolution genes encode mutant gibberellin response
modulators
Unpublished
3 (bases 1 to 1872)
Richards,D.E.
Direct Submission
Submitted (25-MAY-1999) Richards D.E., Molecular Genetics, John
Innes Centre, Colney Lane, NR4 7UJ, UNITED KINGDOM
Location/Qualifiers
1..1872
/organism="Triticum aestivum"
/db_xref="taxon:4565"
/gene="rht-Dla"
1..1872
/gene="rht-Dla"
/note="GAI ortholog"
/codon_start=1
/product="gibberellin response modulator"
/protein_id="CAB5155.1"
/db_xref="GI:5640157"
/translating="MKREYQDAGSGGGGGGSGSSDKMMVSAAGGEEVEYDELLAAL
GYKVRASDMADVAQKLEQEMAGMGVGGAGAPDDSFATHLATDTVHYNPTDLSSWV
ESMLSELNAPPPPLPAPOLNASTSTSTVTCGGYFDLPSPVDSSTSYALRPISPAG
ATAPADLADSVDRPKRMTGGSTSSSSSSSLGGGKRSVVVEAAPVVAANATP
ALPVVYVDQTAGIRLVHALLACAEVQENLSAAELVKQIPLLAASQGGAMRKVAA
YFGEALARRVFRFPQPSLLDAADFADLLHAHFYESCPLYKFAHFTANQAILFAFAG
CRRVHVVDGFKOGMOWPALLQALRALPQGGPPFRLTGVPQPDQTDALQOVGWKLA
QFAHTIRVDFQYRGLVAATLADLEPMLOPEGEDPNEEPEVIAVNSVFEMHRLAOP
GALEKVLGTVAVRPRIVTVVEQANHNSGTFLDRFTESLHYSTMFDSLEGGSSGG
PSEVSGAAAPAAAGTDQWVSEYILGRQICNVVACGAERHTERLTGQWRNRKIGNA
GFETVHLGSNAYKQASTLLALFAGGDGKVEEKGECUWHTPLIATSAWRLAGP"
BASE COUNT      265 a      700 c      641 g      266 t
ORIGIN

Query Match      100.0%; Score 51; DB 8; Length 1872;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacagctgctggcgccgctcggtgacaaagtgccgctccgacatggcg 51
|||||
Db 112 GACGAGCTGCTGGCGCGCCTCGGTACAAAGTGCGCGCTCCGACATGGCG 162

RESULT 4
LOCUS      AX005805      2125 bp      DNA      PAT      24-AUG-2000
DEFINITION      Sequence 14 from Patent WO9909174.
ACCESSION      AX005805
VERSION      AX005805.1      GI:9928802
KEYWORDS      bread wheat.
SOURCE      Triticum aestivum
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 2125)
AUTHORS      Harberd,N.P. and Peng,J.
TITLE      Genetic control of plant growth and development
JOURNAL      Patent: WO 9909174-A 14 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)

```

FEATURES	Location/Qualifiers	misc_feature	/note="n is any nucleotide"
source	1..2125		
	/organism="Triticum aestivum"		
	/db_xref="taxon:4565"		
BASE COUNT	325 a 768 c 723 g 309 t		
ORIGIN			
Query Match	100.0%; Score 51; DB 6; Length 2125;		
Best Local Similarity	100.0%; Pred. No. 4.8e-05;		
Matches	51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 gacgagctgctggcgctgggtacaggtgcgcctccgacatggcg 51		
Db	139 GACGAGCTGCTGGCGGCTCGGGTACAAGTGCAGGCTCGGACATGGCG 189		
RESULT 5			
AX005794	2709 bp DNA PAT 24-AUG-2000		
LOCUS	Sequence 3 from Patent WO9909174.		
DEFINITION	AX005794		
ACCESSION	AX005794		
VERSION	AX005794.1 GI:9928799		
KEYWORDS	bread wheat.		
SOURCE	Triticum aestivum		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.		
REFERENCE	1 (bases 1 to 2709)		
AUTHORS	Harberd,N.P. and Peng,J.		
TITLE	Genetic control of plant growth and development		
JOURNAL	Patent: WO 9909174-A 3 25-FEB-1999;		
	HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)		
FEATURES	Location/Qualifiers		
source	1..2709		
	/organism="Triticum aestivum"		
	/db_xref="taxon:4565"		
misc_feature	6		
misc_feature	11		
misc_feature	46		
misc_feature	85		
misc_feature	107		
misc_feature	407		
misc_feature	417		
misc_feature	430		
misc_feature	436		
misc_feature	448		
misc_feature	452		
misc_feature	490		
misc_feature	498		
misc_feature	536		
misc_feature	562		
misc_feature	727		
misc_feature	769		
misc_feature	1444		
Query Match	100.0%; Score 51; DB 6; Length 2709;		
Best Local Similarity	100.0%; Pred. No. 4.5e-05;		
Matches	51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 gacgagctgctggcgctgggtacaggtgcgcctccgacatggcg 51		
Db	273 GACGAGCTGCTGGCGGCTCGGGTACAAGTGCAGGCTCGGACATGGCG 323		
RESULT 6			
AX005867/c	324 bp DNA PAT 24-AUG-2000		
LOCUS	Sequence 76 from Patent WO9909174.		
DEFINITION	AX005867		
ACCESSION	AX005867.1 GI:9928862		
VERSION	AX005867.1		
KEYWORDS	bread wheat.		
SOURCE	Triticum aestivum		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.		
REFERENCE	1 (bases 1 to 324)		
AUTHORS	Harberd,N.P. and Peng,J.		
TITLE	Genetic control of plant growth and development		
JOURNAL	Patent: WO 9909174-A 76 25-FEB-1999;		
	HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)		
FEATURES	Location/Qualifiers		
source	1..324		
	/organism="Triticum aestivum"		
	/db_xref="taxon:4565"		
misc_feature	158		
misc_feature	161		
misc_feature	217		
BASE COUNT	39 a 141 c 85 g 56 t		
ORIGIN			
Query Match	100.0%; Score 49; DB 6; Length 324;		
Best Local Similarity	100.0%; Pred. No. 0.00032;		



BAC clone OSJNBb0022E02 is from Oryza sativa chromosome 3  
The orientation of the sequence is from SP6 to T7 end of the BAC clone.  
Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), genscan and Genscan+ (Chris Burge, <http://CCR-081.mit.edu/GENSCan.html>), GeneMarkHM (Mark Borodovsky, <http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer (Mihaela Portea and Steven Salzberg, [contact.mpartea@tigr.org](http://contact.mpartea@tigr.org)), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).  
Location/Qualifiers  
1. .122497  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="3"  
/map="C944"  
/clone="OSJNBb0022E02"  
complement(<1271..>1579)  
/gene="OSJNBb0022E02.11"  
complement(1271..1579)  
/gene="OSJNBb0022E02.11"  
/notes="predicted by fgenesh"  
complement(1271..1579)  
/gene="OSJNBb0022E02.11"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAK50135.1"  
/db\_xref="GI:13937304"  
/translation="MNTKAAPLVSSPFSVLSPAAVDPRTLPAPLLHPGLSVAP  
SCRTRSSGTLTPVKTLALMAKVVDGWSHSVAFQKWAIVLAENLAGSFEVATS  
"  
<1386..>1640  
/gene="OSJNBb0022E02.18"  
1386..1640  
/gene="OSJNBb0022E02.18"  
/notes="predicted by glimmerR"  
1386..1640  
/gene="OSJNBb0022E02.18"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAK50133.1"  
/db\_xref="GI:13937302"  
/translation="MRASVFTGVHSNPLERVQLGATLRPGCGSGGARNVGRGSTAA  
GSDNTENGEDDNGAAVFESITAPAPNGTTGATVREEL"  
complement(join(<5240..5517,5605..5684,6132..6212,  
6314..6383,6539..6613,6693..6898,7177..7426,7506..7551,  
8773..8842,9564..9633,11257..11341,11440..11524,  
11608..11698,12255..>12557))  
/gene="OSJNBb0022E02.2"  
complement(5240..12557)  
/gene="OSJNBb0022E02.2"  
/notes="EST D47656, AU088605 from this gene"  
complement(join(5439..5517,5605..5684,6132..6212,  
6314..6383,6539..6613,6693..6898,7177..7426,7506..7551,  
8773..8842,9564..9633,11257..11341,11440..11524,  
11608..11698,12255..12388))  
/gene="OSJNBb0022E02.2"  
/codon\_start=1

/product="unknown protein"  
/protein\_id="AAK50136.1"  
/db\_xref="GI:13937305"  
/translation="MORRAQWAGVGKTAQAAAAHAALFCFTLLALAKVDGRTATSW  
WIIFIPWLFHGIVARGFRSMPAPSLPHGRHWAPCHSIVAAPLLIAPELLLCILLES  
RYKSKPTDLKIVFLPALLAFVILILADNRMCRALMPGDEESNDIAELLPHFWFA  
INSVFLIATTTLLKLSGDVGALMGDLFINYGIAECFAFLVCTFWPMIHKSNP  
GSASSAAIRYRDWESGLLPSLEDHEQERLCGLDIGHVMKIPLVIFOVLLCMRL  
ETPPSAQYIPIFALPSPLFILOGAGVFLSLARLEKVLVLLRNGSPVSNYLTISKV  
RDCFAPLHGRSRLGLWMSIDEGSKERQARLFTYESTGYNTFCGYPEYVVKMPKRDIA  
EEMVQALAGQSESTIKCTKQERLQNEKVLRCICYEIGEICWMLLPCRHRLCTC  
SKCKKCPICRPVIERMPVYV"  
complement(12328..12420)  
/rpt\_family="(CGGGG)n"  
complement(12558..12580)  
/rpt\_family="(GA)n"  
join(<14168..15023,15131..15274,15344..16687,16818..17803,  
17887..18048,18262..>18366)  
/gene="OSJNBb0022E02.12"  
14168..18366  
/gene="OSJNBb0022E02.12"  
/notes="similar to gag pol polyprotein GB:AAD12997  
GI:4234852 (2ea may)s)"  
join(14168..15023,15131..15274,15344..16687,16818..17803,  
17887..18048,18262..18366)  
/gene="OSJNBb0022E02.12"  
/codon\_start=1  
/product="putative gag-pol polyprotein"  
/protein\_id="AAK50117.1"  
/db\_xref="GI:13937286"  
/translation="MAGFADALRPDKFTGVHFKRWQIRVTLMLTAMKCFWVSTGKPEG  
VLTAQOKQFEATTLFVGCILSVGLDRVLEVYVMTDAKELWDLNLTFGATDASND  
LYTMEGFHDYKADNSVVEQAEHIOTMAKELLKCVLPDKFVAGCIITAKLPSSNRS  
FGTALAKHKKQEYSVEGLIASLAEVEERAKDAASKGDSOSSANYHVKQAQNSKGYK  
AQQTTFKKKKNNPNQDERTCFVCGQVGHARLQPKQKRAAGACPFYRQESC  
YECGGLFRFSLDFCNKNSHIGCSVDDEANVHSLRCHINFLGMSRLSMCLIPKFS  
IVKGSCHSCVQSKOPKPKHAAERNLAPLELLHSDLCBMGVLTKGGRYFMTLID  
DATRCYVYLLTKDEALDYFKTYKAEVENOLDRIKRLSDRGSEFFSENEPDLCEE  
HGIIHERTPPSPESNGIAERNRTLTDLVNMLDTAGLPKAWGALLTSNVLNRY  
PNRKDKTPEIWIWGRKPSLSYLRTWGLAKVNPITKKRLGPKTVDCVFLGIAHHS  
IAYRFLIVKSEVDMHVGTIMESDATFFESFPMDKTHSGSNQSEIIPSSITPPEQ  
TEHTHELSEEDVSEAPRRSKQRTAKSGDDFTVYLDVDPKSISEAVSPADYWK  
EAVRSMDSTLIANGTWEVTERPKYCKPVGCKVFKKLRDPAKSHGLLVHOMQDYAF  
NCELDEIYNDQPDGFVEQEGKCKLLKSLYGLQAQKQWHEKDKTLTSAGAVN  
EADKCYRHGGEGVLLCYVDLIFGTNLVINEVKSFLSQNDFMDKDLGVAVIL  
NKLIRGENGITLLQSHYVEKILNRFYIDSKPSPPTPYDPSLLLRNKNRIARNOLEYS  
QITIGSLMYSATRPDIISFVSKLSRFTSNPGDDHWRALERVMRYLKGTVELGLHTG  
YPALRGYSDSNWISDVDEIKATSGYVFTLGGAVSWRSCKQTILTRSTMEALTD  
TATVAEWLRLDMLPVVEKPTCEKTIKQEIKKLRSYVGLHPNSEKPGSLHEG  
TITKDRQCLEGGGFETHSILEGTHLCEDLCHWSQSMKILGESSRKLTKDLGV"  
complement(18822..18858)  
/rpt\_family="(CAGGG)n"  
complement(19612..19640)  
/rpt\_family="AT\_rich"  
complement(20459..20487)  
/rpt\_family="AT\_rich"  
<21271..>21576  
/gene="OSJNBb0022E02.13"  
21271..21576  
/gene="OSJNBb0022E02.13"  
/notes="predicted by fgenesh"  
21271..21576  
/gene="OSJNBb0022E02.13"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAK50120.1"  
/db\_xref="GI:13937289"  
/translation="MTTRIVLLSEYSDSHDNGRNGAVGVGHGVGAGGRLHGYI  
ACAVILLASTPTAPLHQDELQHVMTAKDCRWQAPAWDVEDYPTSNDAVSPNS"  
complement(22507..22529)  
/rpt\_family="(TA)n"  
complement(22544..22588)

repeat\_region

repeat\_region

mRNA

gene

CDS

repeat\_region

repeat\_region

repeat\_region

mRNA

gene

CDS

repeat\_region

repeat\_region

```
repeat_region /rpt_family="AT-rich"
complement(23124..23217)
/rpt_family="(GGAGAA)n"
complement(join(<23165..23588,23765..23917,25379..>25551))
/gene="OSJNBb0022E02.14"
complement(23165..25551)
/gene="OSJNBb0022E02.14"
/note="predicted by fgenesh"
complement(join(23165..23588,23765..23917,25379..25551))
/gene="OSJNBb0022E02.14"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50123.1"
/db_xref="GI:1393792"
/translation="MAGSGARAVDNGGGVGGPPAPASLPLPPHPSPRRRTPA
SPIGLTPRLRCOYSTGSLPESAAGHAVATPSRPPGADADEGVLGLSSAAAV
WMGAQASGCGENAESMAVALSACPYPAGDIALAKEATHGCGVMKGVHGVFTNSL
VCMYKLGEDMDNAKAFDATERKNIWTVNTLITSYATAGLCDEALDVLAAQMEQIGTV
Query Match 88.6%; Score 45.2; DB 8; Length 122497;
Best Local Similarity 94.0%; Pred. No. 0.00059;
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 gacgagctgctgcgcgcctcggtacaaagtgcgcgctccgcacatggc 50
|||||
DB 69197 GACGAGCTGCTGCGCGCTCGGTACAAGGTGCGGCTCGCATGCGC 69246
|||||

RESULT 10
AX005810 725 bp DNA PAT 24-AUG-2000
LOCUS Sequence 19 from Patent WO9909174.
DEFINITION
ACCESSION AX005810
VERSION AX005810.1 GI:9928807
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 725)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 19 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source Location/Qualifiers
1..725
/organism="Oryza sativa"
/db_xref="taxon:4530"
misc_feature 171
/note="n is any nucleotide"
misc_feature 302
/note="n is any nucleotide"
misc_feature 427
/note="n is any nucleotide"
misc_feature 444
/note="n is any nucleotide"
misc_feature 459
/note="n is any nucleotide"
misc_feature 711
/note="n is any nucleotide"
BASE COUNT 98 a 226 c 276 g 119 t 6 others
ORIGIN

Query Match 86.7%; Score 44.2; DB 6; Length 725;
Best Local Similarity 93.9%; Pred. No. 0.0056;
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 acgagctgctgcgcgcctcggtacaaagtgcgcgctccgcacatggc 50
|||||
DB 107 ACGAGCTGCTGCGCGCTCGGTACAAGGTGCGGCTCGCATGCGC 155
|||||
```

```
RESULT 11
AX005860 200 bp DNA PAT 24-AUG-2000
LOCUS Sequence 69 from Patent WO9909174.
DEFINITION
ACCESSION AX005860
VERSION AX005860.1 GI:9928855
KEYWORDS bread wheat.
SOURCE Triticum aestivum
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidaeae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 200)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 19 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source Location/Qualifiers
1..200
/organism="Triticum aestivum"
/db_xref="taxon:4565"
misc_feature 37
/note="n is any nucleotide"
misc_feature 66
/note="n is any nucleotide"
misc_feature 164
/note="n is any nucleotide"
misc_feature 172
/note="n is any nucleotide"
misc_feature 179
/note="n is any nucleotide"
misc_feature 185
/note="n is any nucleotide"
misc_feature 188
/note="n is any nucleotide"
BASE COUNT 24 a 58 c 84 g 27 t 7 others
ORIGIN

Query Match 84.7%; Score 43.2; DB 6; Length 200;
Best Local Similarity 93.8%; Pred. No. 0.016;
Matches 45; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 gacgagctgctgcgcgcctcggtacaaagtgcgcgctccgcacatg 48
|||||
DB 79 GACTATCTGCTGCGCGCTCGGTACAAGGTGCGGCTCGACAGG 126
|||||

RESULT 12
AX005807 302 bp DNA PAT 24-AUG-2000
LOCUS Sequence 16 from Patent WO9909174.
DEFINITION
ACCESSION AX005807
VERSION AX005807.1 GI:9928804
KEYWORDS Zea mays.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 302)
AUTHORS Harberd,N.P. and Peng,J.
TITLE Genetic control of plant growth and development
JOURNAL Patent: WO 9909174-A 16 25-FEB-1999;
HARBERD NICHOLAS PAUL (GB); PENG JINRONG (GB)
FEATURES
source Location/Qualifiers
1..302
/organism="Zea mays"
/db_xref="taxon:4577"
BASE COUNT 52 a 82 c 123 g 45 t
ORIGIN
```





---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 20:52:13 ; Search time 142.61 Seconds  
(without alignments)  
306.596 Million cell updates/sec

Title: US-09-485-529-105

Perfect score: 51

Sequence: 1 gacgagctgctggcgcgct.....tcgcgcctccgacatggcg 51

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	51	20	AA336284
2	51	100.0	453	20	AA336263
3	51	100.0	2125	20	AA336279
4	51	100.0	2709	20	AA336275
5	49	96.1	324	20	AA336273
6	45.2	86.6	770	20	AA336277
7	44.2	86.7	725	20	AA336276
8	43.2	84.7	200	20	AA336266
9	39.8	78.0	302	20	AA336281
10	39.8	78.0	371	20	AA336282
11	39.8	78.0	2255	20	AA336280

12	39	76.5	416	20	AA336283	Partial sequence o
13	29.8	58.4	1542	21	AA360717	DNA encoding a bet
14	29.8	58.4	15079	16	AAQ91580	S. clavuligerus cl
15	29.4	57.6	309	20	AA336261	DNA sequence obtai
16	28.6	56.1	230	20	AA336267	DNA sequence obtai
17	27.6	54.1	1602	21	AA345745	Arabidopsis thalia
18	27.4	53.7	285	20	AA336271	DNA sequence obtai
19	26	51.0	457	21	AA365312	Arabidopsis SCL ES
20	26	51.0	1951	22	AA365312	A. thaliana transc
21	26	51.0	1951	22	AA365312	Arabidopsis thalia
22	26	51.0	1964	18	AA365312	Arabidopsis thalia
23	25.8	50.6	773	21	AA365312	Maize glutathione-
24	25.8	50.6	1100	21	AA365312	Maize glutathione-
25	25.4	49.8	1764	22	AA365312	A. thaliana transc
26	25.4	49.8	1764	22	AA365312	Arabidopsis thalia
27	25	49.0	642	20	AA334188	Mycobacterium spec
28	25	49.0	774	20	AA334189	Mycobacterium spec
29	24.6	48.2	12381	21	AA358381	Streptomyces averm
30	24.6	48.2	30690	21	AA358381	S. avermitilis ave
31	24.4	47.8	1779	22	AA325480	Nucleotide sequenc
32	24.4	47.8	1779	22	AA325481	Nucleotide sequenc
33	24.4	47.8	4838	21	AA359353	DNA encoding a car
34	24.4	47.8	9048	22	AA390812	B. lactofermentum
35	24	47.1	4428	22	AA390812	B. lactofermentum
36	24	47.1	4428	22	AA390812	Bovine alpha(III)
37	24	47.1	22081	22	AA390812	Bovine alpha(III)
38	23.8	46.7	2870	16	AA391177	Human neuroblastom
39	23.4	45.9	1866	12	AA391177	Lipase and accesso
40	23.4	45.9	109519	22	AA391177	Human RXR-alpha co
41	23	45.1	632	22	AA391177	Micromonospora DNA
42	23	45.1	1458	19	AA391177	Murine Hoxcl3 prot
43	23	45.1	1590	22	AA391177	Keratan sulphate 6
44	23	45.1	2451	22	AA391177	Pseudomonas sp lip
45	22.8	44.7	529	22	AA391177	Murine GC13 fragme
						Human cDNA clone (

#### ALIGNMENTS

RESULT 1

AA336284  
ID AAX36284 standard; DNA; 51 BP.

XX AC AAX36284;

XX DT 16-JUL-1999 (first entry)

XX DE Oligonucleotide derived from the wheat rht gene.

XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;

XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;

XX KW paclobutrazol; ss.

XX OS Triticum aestivum.

XX PN WO9909174-A1.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-GB02383.

XX PR 13-AUG-1997; 97GB-0017192.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Harberd NP, Peng J, Richards DE;

XX DR WPI; 1999-181040/15.

XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which

XX PT provides inhibition of the growth of plants, which inhibition is

XX PT antagonised by gibberellin, used to confer a dwarf phenotype



Query Match 100.0%; Score 51; DB 20; Length 2125;  
Best Local Similarity 100.0%; Pred. No. 6.5e-08;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacgagctgctggcgctcggtacaaagtgcgcgctcccgacatggcg 51  
|||||  
Db 139 gacgagctgctggcgctcggtacaaagtgcgcgctcccgacatggcg 189

RESULT 4  
AA36275  
ID AAX36275 standard; DNA; 2709 BP.  
XX  
AC AAX36275;  
XX  
DT 16-JUL-1999 (first entry)  
XX  
DE Composite DNA sequence of wheat Rht gene.  
XX  
DE Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.  
XX  
OS Triticum aestivum.  
XX  
PN WO9909174-A1.  
XX  
PD 25-FEB-1999.  
XX  
PF 07-AUG-1998; 98WO-GB02383.  
XX  
PR 13-AUG-1997; 97GB-0017192.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Harberd NP, Peng J, Richards DE;  
XX  
DR WPI; 1999-181040/15.  
XX  
PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype  
XX  
PS Disclosure; Fig 3a; 88pp; English.  
XX  
CC The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence represents  
CC the composite DNA sequence of wheat Rht gene.  
XX  
SQ Sequence 2709 BP; 473 A; 900 C; 823 G; 482 T; 31 other;

Query Match 100.0%; Score 51; DB 20; Length 2709;  
Best Local Similarity 100.0%; Pred. No. 6.5e-08;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacgagctgctggcgctcggtacaaagtgcgcgctcccgacatggcg 51  
|||||  
Db 273 gacgagctgctggcgctcggtacaaagtgcgcgctcccgacatggcg 323

## RESULT 5

AA36273/C  
ID AAX36273 standard; DNA; 324 BP.  
XX

AC AAX36273;

XX 16-JUL-1999 (first entry)

XX DNA sequence obtained after sequencing wheat Rht clone 5a1.

XX Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.

XX Triticum aestivum.

XX WO9909174-A1.

XX 25-FEB-1999.

XX 07-AUG-1998; 98WO-GB02383.

XX 13-AUG-1997; 97GB-0017192.

XX (PLAN-) PLANT BIOSCIENCE LTD.

XX Harberd NP, Peng J, Richards DE;

XX WPI; 1999-181040/15.

XX New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype  
XX  
PS Disclosure; Fig 2c(6); 88pp; English.

XX The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence was  
CC obtained after partially sequencing wheat Rht clone 5a1.

XX Sequence 324 BP; 39 A; 141 C; 85 G; 56 T; 3 other;

Query Match 96.1%; Score 49; DB 20; Length 324;  
Best Local Similarity 96.1%; Pred. No. 2.7e-07;  
Matches 49; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gacgagctgctggcgctcggtacaaagtgcgcgctcccgacatggcg 51  
|||||  
Db 162 GNCNGCTGCTGGCGGCGCTCGGGTACAAAGGTGGCGGCTCCGACATGGCG 112

## RESULT 6

AA36277  
ID AAX36277 standard; DNA; 770 BP.  
XX

AC AAX36277;

XX 16-JUL-1999 (first entry)

XX Rice EST D39460 sequence, homologous to wheat Rht gene.  
DE  
XX

KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; rice; expressed sequence tag; EST; ss.  
 OS Oryza sativa.  
 XX WO9909174-A1.  
 PN 25-FEB-1999.  
 XX 07-AUG-1998; 98WO-GB02383.  
 PF 13-AUG-1997; 97GB-0017192.  
 XX (PLAN-) PLANT BIOSCIENCE LTD.  
 PA Harberd NP, Peng J, Richards DE;  
 PI WPI; 1999-181040/15.  
 XX P-PSDB; AAY02538.  
 DR New Triticum Aestivum polynucleotides - encode a polypeptide which  
 XX provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Claim 13; Fig 6a; 88pp; English.  
 PS The specification describes polypeptides encoded by the Rht gene (and  
 XX its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC rice expressed sequence tag (EST) AAD39460, which is homologous to the  
 CC wheat Rht gene.  
 XX Sequence 770 BP; 104 A; 252 C; 294 G; 120 T; 0 other;  
 SQ

Query Match 88.6%; Score 45.2; DB 20; Length 770;  
 Best Local Similarity 94.0%; Pred. No. 4.6e-06;  
 Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gacgagctgctggcgctcggtacaaaggtgcgcgcctccgacatggc 50  
 |||||||  
 DB 114 gacgagctgctggcgctcggtacaaaggtgcgcgcctccgacatggc 163  
 |||||||

RESULT 7  
 AAX36276  
 ID AAX36276 standard; DNA; 725 BP.  
 XX AAX36276;  
 AC 16-JUL-1999 (first entry)  
 DT Partial cDNA sequence of rice D39460 (a wheat Rht gene homologue).  
 XX  
 DE Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 XX antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX Oryza sativa.  
 OS WO9909174-A1.  
 PN 25-FEB-1999.  
 XX 07-AUG-1998; 98WO-GB02383.  
 PF 13-AUG-1997; 97GB-0017192.  
 XX (PLAN-) PLANT BIOSCIENCE LTD.  
 PA Harberd NP, Peng J, Richards DE;  
 PI WPI; 1999-181040/15.  
 XX P-PSDB; AAY02538.  
 DR New Triticum Aestivum polynucleotides - encode a polypeptide which  
 XX provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Claim 13; Fig 6a; 88pp; English.  
 PS The specification describes polypeptides encoded by the Rht gene (and  
 XX its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC rice expressed sequence tag (EST) AAD39460, which is homologous to the  
 CC wheat Rht gene.  
 XX Sequence 770 BP; 104 A; 252 C; 294 G; 120 T; 0 other;  
 SQ

PD 25-FEB-1999.  
 XX 07-AUG-1998; 98WO-GB02383.  
 XX 13-AUG-1997; 97GB-0017192.  
 PR (PLAN-) PLANT BIOSCIENCE LTD.  
 XX Harberd NP, Peng J, Richards DE;  
 XX WPI; 1999-181040/15.  
 DR New Triticum Aestivum polynucleotides - encode a polypeptide which  
 XX provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Disclosure; Fig 4a; 88pp; English.  
 PS The specification describes polypeptides encoded by the Rht gene (and  
 XX its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the partial cDNA sequence of rice AAD39460 (a wheat Rht gene homologue).  
 XX Sequence 725 BP; 98 A; 226 C; 276 G; 119 T; 6 other;  
 SQ

Query Match 86.7%; Score 44.2; DB 20; Length 725;  
 Best Local Similarity 93.9%; Pred. No. 9.7e-06;  
 Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 acgagctgctggcgctcggtacaaaggtgcgcgcctccgacatggc 50  
 |||||||  
 DB 107 acgagctgctggcgctcggtacaaaggtgcgcgcctccgacatggc 155  
 |||||||

RESULT 8  
 AAX36266  
 ID AAX36266 standard; DNA; 200 BP.  
 XX AAX36266;  
 AC 16-JUL-1999 (first entry)  
 DT DNA sequence obtained after sequencing wheat Rht clone 14a1.  
 XX  
 DE Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX Triticum aestivum.  
 OS WO9909174-A1.  
 PN 25-FEB-1999.  
 XX 07-AUG-1998; 98WO-GB02383.  
 PF 13-AUG-1997; 97GB-0017192.  
 XX (PLAN-) PLANT BIOSCIENCE LTD.  
 PA Harberd NP, Peng J, Richards DE;  
 PI WPI; 1999-181040/15.  
 XX P-PSDB; AAY02538.  
 DR New Triticum Aestivum polynucleotides - encode a polypeptide which  
 XX provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Disclosure; Fig 4a; 88pp; English.  
 PS The specification describes polypeptides encoded by the Rht gene (and  
 XX its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the partial cDNA sequence of rice AAD39460 (a wheat Rht gene homologue).  
 XX Sequence 725 BP; 98 A; 226 C; 276 G; 119 T; 6 other;  
 SQ









Search completed: December 19, 2001, 22:25:14  
Job time: 5581 sec

Query Match 58.4%; Score 29.8; DB 16; Length 15079;  
Best Local Similarity 75.5%; Pred. No. 0.47;  
Matches 37; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 3 cgagctgtcggcggtcggtacaaagtcgcgcctccacatggcg 51  
|||||  
Db 4848 cgagctgtcggcggtcgcgcctccacatggcg 4896  
|||||

RESULT 15  
AAX36261  
ID AAX36261 standard; DNA; 309 BP.  
XX  
AC AAX36261;  
XX  
DT 16-JUL-1999 (first entry)  
XX  
DE DNA sequence obtained after sequencing wheat Rht clone 14a1.  
XX  
KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.  
XX  
OS Triticum aestivum.  
XX  
PN WO9909174-A1.  
XX  
PD 25-FEB-1999.  
XX  
PF 07-AUG-1998; 98WO-GB02383.  
XX  
PR 13-AUG-1997; 97GB-0017192.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Harberd NP, Peng J, Richards DE;  
XX  
WPI; 1999-181040/15.  
XX  
PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype  
XX  
PS Disclosure; Fig 2b(7); 88pp; English.  
XX  
CC The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence was  
CC obtained after partially sequencing wheat Rht clone 14a1.  
XX  
SQ Sequence 309 BP; 47 A; 102 C; 102 G; 45 T; 13 other;

Query Match 57.6%; Score 29.4; DB 20; Length 309;  
Best Local Similarity 93.8%; Pred. No. 0.55;  
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 20 tcgggtacaaagtcgcgcctccacatggcg 51  
|  
Db 16 ttgggtacaaagtcgcgcctccacatggcg 47  
|





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 16:58:38 ; Search time 1383.43 Seconds  
(without alignments)  
396.142 Million cell updates

Title: US-09-485-529-105  
Perfect score: 51  
Sequence: 1 qacqagctgctgqgc

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

C	41	RESULT	1
	42	D39460	
	43	LOCUS	
	44	DEFINITION	
	45	ACCESSION	
		VERSION	
		KEYWORDS	
		SOURCE	
		ORGANISM	
		REFERENCE	

**pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	45.2	88.6	263	11	D39460	R1CS0803A R
2	45.2	88.6	388	11	C27475	C27475 R1CS0803A R
3	45.2	88.6	399	10	A0091413	A0091413 A0091413
4	45.2	88.6	467	10	A0222715	A0222715 A0222715
5	44.6	87.5	844	11	BF268018	BF268018 HV_CEA001
6	40.4	79.2	711	10	BE659955	BE659955 1077 Gmax
7	36.4	71.4	455	10	BE659954	BE659954 746 GmaxS
8	35.6	69.8	355	10	AV410222	AV410222 AV410222
9	35.6	69.8	484	10	AV422153	AV422153 AV422153
10	33.4	65.5	365	10	AV409731	AV409731 AV409731
11	33.4	65.5	478	10	AW720697	AW720697 L1NEST7h8
12	33.4	65.5	586	11	B1419686	B1419686 L1NEST47e

Query Match 88.6% Score 45.2 DB 11 Length 263

## ALIGNMENTS

RESULT	1
D39460	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

**pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	45.2	88.6	263	11	D39460	RICS0803A R
2	45.2	88.6	388	11	C27475	C27475 RICE
3	45.2	88.6	399	10	A0091413	A0091413 A0091413
4	45.2	88.6	467	10	A0222715	A0222715 A0222715
5	44.6	87.5	844	11	BF268018	HV_CEA001
6	40.4	79.2	711	10	BE659955	BE659955 1077 Gmax
7	36.4	71.4	455	10	BE659954	746 GmaxS
8	35.6	69.8	355	10	AV410222	AV410222 AV410222
9	35.6	69.8	484	10	AV422153	AV422153 AV422153
10	33.4	65.5	365	10	AV409731	AV409731 AV409731
11	33.4	65.5	478	10	AW720697	AW720697 L1NEST7h8
12	33.4	65.5	586	11	B1419686	B1419686 L1NEST47e

Query Match 88.6% Score 45.2; DB 11; Length 263;

Best Local Similarity 94.0%; Pred. NO. 0.00071;  
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gacgagctgctgcgcgcgtcggtacaaaggtgcgcgcctccgacatggc 50  
|||||  
Db 97 GACGAGCTGCTGCGCGCTCGGGTACAAAGGTGCGGTCTCGCGACATGGC 146

## RESULT 2

LOCUS C27475 388 bp mRNA EST 06-AUG-1997  
DEFINITION C27475 Rice callus cDNA Oryza sativa cDNA clone C51976\_1A, mRNA sequence.  
ACCESSION C27475  
VERSION C27475.1 GI:2311320  
KEYWORDS EST.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 388)

REFERENCE Yamamoto, K. and Sasaki, T.  
AUTHORS Rice cDNA from callus 1997  
TITLE Unpublished (1997)  
JOURNAL  
COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/.

## FEATURES

source  
1..388  
Location/Qualifiers  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="C51976\_1A"  
/clone\_lib="Rice callus cDNA"  
/tissue\_type="callus"  
/dev\_stage="callus"  
BASE COUNT 83 a 102 c 134 g 65 t 4 others  
ORIGIN

Query Match 88.6%; Score 45.2; DB 11; Length 388;  
Best Local Similarity 94.0%; Pred. NO. 0.00072;  
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gacgagctgctgcgcgcgtcggtacaaaggtgcgcgcctccgacatggc 50  
|||||  
Db 271 GACGAGCTGCTGCGCGCTCGGGTACAAAGGTGCGGTCTCGCGACATGGC 320

## RESULT 3

LOCUS AU091413 399 bp mRNA EST 05-JUN-2000  
DEFINITION AU091413 Rice cDNA from immature leaf including apical meristem sequence.  
ACCESSION AU091413  
VERSION AU091413.1 GI:8251089  
KEYWORDS EST.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 399)

REFERENCE Sasaki, T. and Yamamoto, K.  
AUTHORS Rice cDNA from immature leaf including apical meristem (2000)  
TITLE Unpublished (2000)  
JOURNAL  
COMMENT Contact: Takuji Sasaki

National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = 'RGP'.  
E60220\_1A.

## FEATURES

source  
1..399  
Location/Qualifiers  
/organism="Oryza sativa"  
/strain="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="E60220"  
/clone\_lib="Rice cDNA from immature leaf including apical meristem (under short day condition)"  
/dev\_stage="immature leaf including apical meristem (under short day condition)"  
BASE COUNT 81 a 122 c 117 g 79 t  
ORIGIN

Query Match 88.6%; Score 45.2; DB 10; Length 399;  
Best Local Similarity 94.0%; Pred. NO. 0.00073;  
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gacgagctgctgcgcgcgtcggtacaaaggtgcgcgcctccgacatggc 50  
|||||  
Db 346 GACGAGCTGCTGCGCGCTCGGGTACAAAGGTGCGGTCTCGCGACATGGC 395

## RESULT 4

LOCUS AU222715 467 bp mRNA EST 30-JUL-2001  
DEFINITION AU222715 Rice shoot Oryza sativa cDNA clone S0583, mRNA sequence.  
ACCESSION AU222715  
VERSION AU222715.1 GI:15008327  
KEYWORDS EST.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 467)

REFERENCE Sasaki, T. and Yamamoto, K.  
AUTHORS Rice cDNA from etiolated shoot (2001)  
TITLE Unpublished (2001)  
JOURNAL  
COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = 'RGP'.  
S0583\_97A.

## FEATURES

source  
1..467  
Location/Qualifiers  
/organism="Oryza sativa"  
/strain="Nipponbare, sub\_species Japonica"  
/db\_xref="taxon:4530"  
/clone="S0583"  
/clone\_lib="Rice shoot"  
/note="Etiolated shoot (8 days old)"  
BASE COUNT 90 a 136 c 176 g 65 t  
ORIGIN

Query Match 88.6%; Score 45.2; DB 10; Length 467;  
Best Local Similarity 94.0%; Pred. NO. 0.00073;  
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gacgagctgctgcgcgcgtcggtacaaaggtgcgcgcctccgacatggc 50

```
|||||
Db 198 GACGAGTCTGTCGGCGCTCGGCTACAGGTGCGTCCGACATGGC 247
|||||

RESULT 5
BF268018 844 bp mRNA 09-MAR-2001
LOCUS HV_CEA0019J17f Hordeum vulgare seedling green leaf EST library
DEFINITION HVCNA0004 (Erysiphe infected & control) Hordeum vulgare cDNA clone
ACCESSION BF268018
VERSION BF268018.2 GI:13263734
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 844)
AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
Wood, T.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics
JOURNAL Unpublished (2000)
COMMENT On Nov 17, 2000 this sequence version replaced gi:11199013.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: ANTATACCTCACTAAGGG
High quality sequence stop: 587.
Location/Qualifiers
1. .844
/organism="Hordeum vulgare"
/cultivar="CI16155 (Mia13)"
/db_xref="taxon:4513"
/clone="HV_CEA0019J17f"
/library="HVCNA0004 (Erysiphe infected & control)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"
/notes="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT 150 a 290 c 276 g 128 t
ORIGIN

Query Match 87.5%; Score 44.6; DB 11; Length 844;
Best Local Similarity 92.2%; Pred. No. 0.0011;
Matches 47; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gacgagctgtgcgcgcgtcggtacaaagtgctgcgcctccacatggcg 51
|||||
Db 260 GACGAGTCTGTCGGCGCTCGGCTACAGGTGCGGCGTCCGACATGGCG 310
|||||

RESULT 6
BF659955 711 bp mRNA EST 06-SEP-2000
LOCUS 1077 GmaxSC Glycine max cDNA, mRNA sequence.
DEFINITION BE659955
ACCESSION BE659955
VERSION BE659955.1 GI:9985949
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 455)
Harris, N., Chapman, B.P. and Gijzen, M.
Gene expression in developing soybean seed coats
Unpublished (2000)
Contact: Gijzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@agr.ca.
Location/Qualifiers
1. .455

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 711)
Harris, N., Chapman, B.P. and Gijzen, M.
Gene expression in developing soybean seed coats
Unpublished (2000)
Contact: Gijzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@agr.ca.
Location/Qualifiers
1. .711
/organism="Glycine max"
/cultivar="Harosoy 63"
/db_xref="taxon:3847"
/clone_lib="GmaxSC"
/tissue_type="Seed coats"
/lab_host="E. coli strain XL0LR"
/notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from green seed coats in mid to late developmental stage
, average fresh weight 250 mg per seed. Traces of pod and
embryo tissue also present. Complementary DNA was
synthesized from mRNA using an XhoI-poly(dT)
linker-primer. EcoRI adapters were ligated to the
blunt-ended cDNA fragments and the products were digested
with XhoI for directional cloning into lambda ZAP Express
vector. This lambda library was amplified once using E.
coli host strain XL1 Blue MRF'. Inserts were then
subcloned by mass excision using ExAssist helper phage for
conversion into phagemid vector pBK-CMV in E. coli host
strain XL0LR."
BASE COUNT 186 a 207 c 179 g 134 t 5 others
ORIGIN

Query Match 79.28; Score 40.4; DB 10; Length 711;
Best Local Similarity 88.08; Pred. No. 0.019;
Matches 44; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 gacgagctgtgcgcgcgtcggtacaaagtgctgcgcctccacatggc 50
|||||
Db 174 GACGAGTCTGTCGGCGCTCGGCTACAGGTGCTGCTCCGACATGGC 223
|||||

RESULT 7
BE659954 455 bp mRNA EST 06-SEP-2000
LOCUS 746 GmaxSC Glycine max cDNA, mRNA sequence.
DEFINITION BE659954
ACCESSION BE659954.1 GI:9985948
VERSION BE659954.1
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 455)
Harris, N., Chapman, B.P. and Gijzen, M.
Gene expression in developing soybean seed coats
Unpublished (2000)
Contact: Gijzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@agr.ca.
Location/Qualifiers
1. .455

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 455)
Harris, N., Chapman, B.P. and Gijzen, M.
Gene expression in developing soybean seed coats
Unpublished (2000)
Contact: Gijzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@agr.ca.
Location/Qualifiers
1. .455
```

Query Match 69.8%; Score 35.6; DB 10; Length 355;  
Best Local Similarity 82.0%; Pred. No. 0.44;











GenCore version 4.5  
Copyright (C) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2001, 18:34:13 ; Search time 69.5 Seconds  
(without alignments)  
166.192 Million cell updates/sec

Title: US-09-485-529-105  
Perfect score: 51  
Sequence: 1 gacgagctgctggcgcgct.....tcgcgctcgcgacatggcg 51

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29.8	58.4	1542	4	US-09-385-028-16
2	29.8	58.4	11604	4	US-09-385-028-13
3	29.8	58.4	15079	4	US-09-385-028-1
4	26	51.0	1964	4	US-09-117-853-1
5	25.8	50.6	773	3	US-09-248-335-39
6	25.8	50.6	1100	3	US-09-248-335-53
7	25	49.0	4403765	4	US-09-103-840A-2
8	25	49.0	4411529	4	US-09-103-840A-1
9	24.4	47.8	812	4	US-08-998-416-401
10	24.4	47.8	4837	4	US-09-629-616-1
11	24	47.1	2005	1	US-07-916-901-5
12	24	47.1	4403765	4	US-09-103-840A-2
13	23.4	45.9	1866	1	US-08-336-408B-1
14	23.4	45.9	1866	5	PCT-US91-00399-1
15	22.8	44.7	1236	1	US-08-440-856A-1
16	22.8	44.7	8540	5	PCT-US96-05611A-12
17	22.8	44.7	15378	3	US-08-785-420-1
18	22.8	44.7	36519	3	US-08-923-137-2
19	22.6	44.3	2088	1	US-08-331-394-1
20	22.6	44.3	2088	1	US-08-250-858-1
21	22.6	44.3	2088	1	US-08-446-915-1
22	22.6	44.3	2088	2	US-08-744-139-1
23	22.6	44.3	2088	5	PCT-US95-06639-1
24	22.4	43.9	298	1	US-08-087-772A-7
25	22.4	43.9	1164	1	US-08-087-772A-3
26	22.4	43.9	1360	1	US-08-087-772A-4
27	22.4	43.9	1920	1	US-08-087-772A-1

Sequence 3, Appli  
Sequence 1, Appli  
Sequence 23, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 17, Appli  
Patent No. 5229279  
Patent No. 5512669  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 4, Appli  
Sequence 1, Appli  
US-08-450-962-3  
US-09-103-840A-1  
US-08-241-943-23  
US-08-254-357-1  
US-09-061-702-3  
US-09-046-992-1  
US-08-510-646B-17  
5229279-5  
5229279-5  
5512669-5  
US-08-743-637B-20  
US-08-526-840B-20  
US-08-748-170A-3  
US-09-047-148-1  
US-08-006-676B-2  
US-08-282-845-1  
US-08-428-414A-4  
PCT-US94-00324-2  
US-08-472-358-1

## ALIGNMENTS

RESULT 1  
US-09-385-028-16  
; Sequence 16, Application US/09385028  
; Patent No. 6232106  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paradkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6232106  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
; STREET: The Jenifer Bulding, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/385,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/P57452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1542 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-385-028-16

Query Match 58.4% Score 29.8; DB 4; Length 1542;

```

; APPLICANT: Ashish S. Paradkar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
; Patent No. 6232106
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN,PPLC
; STREET: The Jenner Buliding, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 1418/P57452US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202 638-6666
; TELEFAX: (202) 39305350
; TELEX: RCA 248593 IDEA UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces clavuligerus
;
; US-09-385-028-1
;
; Query Match 58.4%; Score 29.8; DB 4; Length 15079;
; Best Local Similarity 75.5%; Pred.No.0.24; 12; Indels 0;
; Matches 37; Conservative 0; Mismatches 0;
;
; QY 3 cgaagctgcggcgcgctgggtacaaagtcgcgcctccgcacatggcg 51
; |||||
; DB 4848 CGAGTGTCTGGCGCAGCTCCGCTACGCGGTGTGGGCTCCCGATCGGTG 4896
; |||||
;
; RESULT 4
; US-09-117-853-1
; Sequence 1, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0

```

Query Match 50.6%; Score 25.8; DB 3; Length 1100;  
Best Local Similarity 73.3%; Pred. No. 4;  
Matches 33; Conservative 0; Mismatches 12; Indels 0

**QY** 1 gacgagctgctggcggcgctcgggtacaaggatgcgcgccctccacatgg 49  
| | | | | | | | | | | | | | | | | | | |  
**Pb** 207313 gtcatattctgttcgcgcctccgaatcaaggatgccaacaccactcgg 207361  
| | | | | | | | | | | | | | | | | | | |

```

; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEBACTERIUM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: OP945CIP
; CURRENT APPLICATION NUMBER: US/09/629,616
; CURRENT FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4837
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (283)..(1461)
; NAME/KEY: CDS
; LOCATION: (1470)..(4808)
; US-09-629-616-1

Query Match 47.8%; Score 24.4; DB 4; Length 4837;
Best Local Similarity 73.8%; Pred. No. 11;
Matches 31; Conservative 0; Mismatches 11; Indels 0;

QY 3 cgaagctcgtggcgcgctcggtggtacaaaggtgcgcgcctccga 44
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 4445 ccagcgctggcgtcgatgggctacaagatcctcgccaccga 4486

RESULT 11
US-07-916-901-5/c
; Sequence 5, Application US/07916901
; Patent No. 5364772
; GENERAL INFORMATION:
; APPLICANT: Granneman, James G.
; APPLICANT: Lahners, Kristine N.
; APPLICANT: Rao, Donald D.
; TITLE OF INVENTION: @ @3-ADRENERGIC RECEPTOR PROTEIN AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REISING, ETHINGTON, BARNARD, PERRY &
; ADDRESSEE: MILTON
; STREET: 201 W. Big Beaver - Ste. 400; P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/916,901
; FILING DATE: 19920720
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohr, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-324(WSU)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (313) 689-3554
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2005 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..1250

```

FILING DATE: 08-NOV-1994









OM of: US-09-485-529-104 to: N\_Geneseq\_1101.\* out\_format : pfs

Date: Dec 19, 2001 11:20 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-o=/cgn2.1/USPTO\_spool/US09485529/runat.19122001.165005.14373/app\_query.fasta.1.77  
-DB=N\_Geneseq\_1101 -QPMF=fastap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR\_SCORE=epct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09485529 -CGN1\_1\_248 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
-WAIT -THREADS=1

Search information block:

Query: US-09-485-529-104

Query length: 17

Database: N\_Geneseq\_1101.\*

Database sequences: 930621

Database length: 428662619

Search time (sec): 145.180000

score.list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336284 +	81.00	306.70	5.2e-09	51	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336283 +	81.00	285.09	8.4e-08	453	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336279 +	81.00	269.80	6.0e-07	2125	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336275 +	81.00	267.40	8.1e-07	2709	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336281 +	78.00	277.99	2.1e-07	302	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336282 +	78.00	275.96	2.7e-07	371	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336277 +	78.00	268.73	6.8e-07	770	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336280 +	78.00	258.10	2.7e-06	2255	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336276 +	72.00	247.12	1.1e-05	725	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336273 +	70.00	247.68	1.0e-05	324	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336272 +	70.00	244.28	1.6e-05	457	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336271 +	70.00	231.87	7.7e-05	1602	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336270 +	70.00	230.92	8.7e-05	1764	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336268 +	70.00	230.92	8.7e-05	1764	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336267 +	70.00	229.92	9.9e-05	1951	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336266 +	70.00	229.92	9.9e-05	1951	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336265 +	70.00	229.86	0.00010	1764	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336264 +	66.00	216.03	0.0006	1779	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336263 +	66.00	216.03	0.0006	1779	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336262 +	65.00	233.95	5.9e-05	200	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336261 +	62.00	215.59	0.0006	416	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336260 +	46.00	159.30	0.8500	309	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336259 +	45.00	145.98	4.4e-05	817	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336258 +	45.00	143.12	6.77	1091	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336257 +	44.00	136.55	13.73	1458	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336256 +	44.00	125.68	63.41	4374	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336255 +	44.00	113.21	313.78	15424	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336254 +	44.00	113.21	313.78	15424	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336253 +	44.00	113.20	314.45	15450	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336252 +	43.00	171.61	0.1755	29	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336251 +	42.00	134.32	20.93	864	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336250 +	42.00	134.25	21.12	870	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336249 +	42.00	133.99	21.86	894	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336248 +	42.00	133.99	21.86	894	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336247 +	42.00	130.09	36.01	1325	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336246 +	42.00	129.11	40.87	1464	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336245 +	42.00	128.52	44.09	1564	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336244 +	42.00	128.45	44.45	1564	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336243 +	41.00	143.50	6.45	235	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336242 +	41.00	143.50	6.45	235	!

/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336241 +	41.00	143.50	6.45	235	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336240 +	41.00	143.50	6.45	235	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336239 +	41.00	141.59	8.24	285	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336238 +	41.00	134.98	19.24	556	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336237 +	41.00	125.32	66.39	1476	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336236 +	41.00	125.03	68.97	1521	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336235 +	41.00	125.03	68.97	1521	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336234 +	41.00	125.03	68.97	1521	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336233 +	41.00	124.48	74.42	1608	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336232 +	41.00	119.24	144.88	2730	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336231 +	41.00	118.68	155.60	2888	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336230 +	41.00	111.65	383.54	5880	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336229 +	41.00	87.15	8.9e+03	6993	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336228 +	41.00	78.23	2.8e+04	1723	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336227 +	41.00	71.22	6.7e+04	3499	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336226 +	41.00	71.22	6.7e+04	3499	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336225 +	41.00	62.60	2.9e+05	8370	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336224 +	41.00	57.24	3.7e+05	1437	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336223 +	40.00	140.01	10.09	230	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336222 +	40.00	139.03	11.45	254	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336221 +	40.00	130.48	34.28	603	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336220 +	40.00	130.48	34.28	603	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336219 +	40.00	121.14	113.58	1550	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336218 +	40.00	113.82	290.14	3246	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336217 +	40.00	112.68	336.12	3645	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336216 +	40.00	111.16	408.19	4248	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336215 +	40.00	107.65	640.26	6057	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336214 +	40.00	96.23	2.8e+03	1922	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336213 +	40.00	88.94	7.1e+03	4013	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336212 +	40.00	88.94	7.1e+03	4013	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336211 +	40.00	70.12	7.7e+04	2692	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336210 +	39.00	130.42	34.51	417	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336209 +	39.00	130.31	35.04	422	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336208 +	39.00	127.72	48.81	548	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336207 +	39.00	125.26	66.96	703	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336206 +	39.00	124.49	73.92	760	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336205 +	39.00	124.33	75.40	772	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336204 +	39.00	123.78	80.90	816	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336203 +	39.00	116.81	197.98	1652	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336202 +	39.00	115.32	239.58	1920	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336201 +	39.00	114.00	283.61	2193	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336200 +	39.00	111.35	398.30	2866	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336199 +	39.00	111.35	398.30	2866	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336198 +	39.00	111.04	414.59	2958	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336197 +	39.00	110.76	429.76	3043	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336196 +	39.00	110.02	472.85	3281	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336195 +	39.00	109.56	501.56	3437	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336194 +	39.00	109.22	523.70	3556	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336193 +	39.00	106.96	699.46	4467	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336192 +	39.00	106.68	725.19	4596	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336191 +	39.00	105.78	814.18	5035	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336190 +	39.00	103.86	1.0e+03	6111	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336189 +	39.00	95.49	3.0e+03	1424	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336188 +	39.00	90.40	5.8e+03	2382	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336187 +	39.00	83.07	1.5e+04	4999	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336186 +	38.00	118.46	160.06	961	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336185 +	38.00	115.78	225.71	1260	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336184 +	38.00	114.12	279.45	1491	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336183 +	38.00	114.12	279.45	1491	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336182 +	38.00	113.77	292.36	1545	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336181 +	38.00	111.04	414.94	2036	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336180 +	38.00	107.15	682.87	3015	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336179 +	38.00	107.15	682.87	3015	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336178 +	38.00	106.90	705.36	3093	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336177 +	38.00	103.92	1.0e+03	4178	!
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AA336176 +	38.00	103.75	1.1e+03	4250	!

/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAAS9383	-	38.00	99.45	1.8e+03	6568	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH68532	-	37.00	56.41	4.1e+05	3499
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI58199	-	38.00	92.06	4.7e+03	13857	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAF07582	-	36.50	112.54	342.29	998
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI51590	-	38.00	91.35	5.2e+03	14896	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ23890	-	36.50	111.22	405.23	1140
/SIDS2/cgcgdata/geneseq/geneseq/NA1996.DAT:AAI16239	-	38.00	91.21	5.3e+03	15101	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ23898	-	36.50	106.79	715.29	1784
/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAQ32002	-	38.00	91.21	5.3e+03	15108	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ23191	-	36.00	122.38	96.84	306
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAV35000	-	38.00	79.51	2.4e+04	49272	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH94307	-	36.00	120.26	127.05	379
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH68526	-	38.00	60.12	2.6e+05	349980	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH99977	-	36.00	119.96	132.17	391
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAAO0364	-	37.00	126.28	58.74	300	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH94526	-	36.00	119.93	132.60	392
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ57184	-	37.00	124.02	78.49	377	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI57275	-	36.00	119.46	140.81	411
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ19527	-	37.00	123.94	79.29	380	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI57287	-	36.00	119.34	142.98	416
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ56531	-	37.00	123.73	81.41	388	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI38794	-	36.00	119.34	142.98	416
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ93729	-	37.00	122.03	101.32	461	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI39081	-	36.00	119.34	142.98	416
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ93441	-	37.00	120.97	116.03	513	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI40323	-	36.00	116.51	205.66	554
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ93265	-	37.00	120.93	116.32	514	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ00971	-	36.00	115.90	222.29	589
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ93687	-	37.00	120.03	130.86	564	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ35532	-	36.00	115.11	246.01	638
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI19269	-	37.00	119.84	134.11	575	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ38632	-	36.00	113.87	288.32	723
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAI44439	-	37.00	119.84	134.11	575	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH05152	-	36.00	113.30	310.24	766
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ93449	-	37.00	118.97	149.98	628	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH05174	-	36.00	113.26	311.79	769
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ53463	-	37.00	118.94	150.59	630	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ46695	-	36.00	112.25	355.09	852
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ41378	-	37.00	117.44	182.49	733	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ34384	-	36.00	112.22	356.68	855
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH04767	-	37.00	117.13	189.78	756	/SIDS2/cgcgdata/geneseq/geneseq/NA1989.DAT:AAZ92407	-	36.00	111.81	375.84	891
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ22282	-	37.00	117.11	190.42	758	/SIDS2/cgcgdata/geneseq/geneseq/NA1996.DAT:AAI10417	-	36.00	111.21	405.52	946
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ22274	-	37.00	117.11	190.42	758	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ00972	-	36.00	110.91	421.90	976
/SIDS2/cgcgdata/geneseq/geneseq/NA1994.DAT:AAZ78087	-	37.00	114.80	255.96	957	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH24102	-	36.00	109.69	493.31	1104
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ54008	-	37.00	114.56	264.14	981	/SIDS2/cgcgdata/geneseq/geneseq/NA1996.DAT:AAZ34181	-	36.00	109.04	536.22	1179
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ53013	-	37.00	114.55	264.48	982	/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ66893	-	36.00	109.04	536.22	1179
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ74768	-	37.00	113.05	320.31	1142	/SIDS2/cgcgdata/geneseq/geneseq/NA1990.DAT:AAQ02853	-	36.00	108.17	599.29	1287
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ02679	-	37.00	112.67	336.41	1187	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ48798	-	36.00	108.11	604.02	1295
/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAQ23531	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ44799	-	36.00	106.82	712.48	1475
/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAQ27201	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ59960	-	36.00	106.32	759.38	1551
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ93791	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ34023	-	36.00	104.80	923.10	1809
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ77314	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ54745	-	36.00	104.80	923.10	1809
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAV58013	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAQ20953	-	36.00	104.78	925.04	1812
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ09719	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA1991.DAT:AAQ11780	-	36.00	104.53	956.25	1860
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ09757	-	37.00	111.20	406.15	1377	/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAQ30028	-	36.00	104.53	956.25	1860
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ85666	-	37.00	109.31	517.61	1667	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ47589	-	36.00	103.12	1.1e+03	2143
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ06376	-	37.00	108.70	559.72	1773	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ36382	-	36.00	103.12	1.1e+03	2143
/SIDS2/cgcgdata/geneseq/geneseq/NA1992.DAT:AAZ06374	-	37.00	107.17	681.26	2070	/SIDS2/cgcgdata/geneseq/geneseq/NA1994.DAT:AAZ64125	-	36.00	103.07	1.2e+03	2154
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ93788	-	37.00	107.60	644.73	1982	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ22684	-	36.00	103.07	1.2e+03	2154
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ77313	-	37.00	107.60	644.73	1982	/SIDS2/cgcgdata/geneseq/geneseq/NA1994.DAT:AAZ64125	-	36.00	103.07	1.2e+03	2154
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ58009	-	37.00	107.60	644.73	1982	/SIDS2/cgcgdata/geneseq/geneseq/NA1994.DAT:AAZ64125	-	36.00	103.05	1.2e+03	2160
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ09754	-	37.00	107.60	644.73	1982	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ19347	-	36.00	103.05	1.2e+03	2160
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ06375	-	37.00	107.18	680.01	2067	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ91707	-	36.00	103.05	1.2e+03	2160
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ06374	-	37.00	107.17	681.26	2070	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ42650	-	36.00	102.31	1.3e+03	2328
/SIDS2/cgcgdata/geneseq/geneseq/NA1995.DAT:AAZ98750	-	37.00	107.02	694.17	2108	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ22322	-	36.00	101.93	1.3e+03	2417
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ57475	-	37.00	106.99	697.17	2108	/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ60327	-	36.00	99.48	1.8e+03	3098
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ67768	-	37.00	106.63	729.64	2185	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH15743	-	36.00	99.14	1.9e+03	3206
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ85833	-	37.00	106.02	789.89	2326	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH15760	-	36.00	99.14	1.9e+03	3206
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ31279	-	37.00	106.02	789.89	2326	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ15760	-	36.00	97.08	2.5e+03	3947
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ31280	-	37.00	105.89	802.41	2355	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ34839	-	36.00	97.07	2.5e+03	3953
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ24305	-	37.00	105.64	829.31	2417	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ03014	-	36.00	97.00	2.5e+03	3980
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ31278	-	37.00	105.63	829.74	2418	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ80156	-	36.00	95.31	3.1e+03	4723
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH17466	-	37.00	103.15	1.1e+03	3109	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH24959	-	36.00	92.94	4.2e+03	6000
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ77132	-	37.00	100.12	1.7e+03	4221	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ19992	-	36.00	90.69	5.6e+03	7528
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ59353	-	37.00	98.72	2.0e+03	4838	/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ52292	-	36.00	88.88	7.1e+03	9048
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ67767	-	37.00	96.47	2.7e+03	6136	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAH81480	-	36.00	84.71	1.2e+04	1378
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ13066	-	37.00	96.22	2.8e+03	6259	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ65348	-	36.00	83.86	1.4e+04	1502
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ57305	-	37.00	93.78	3.8e+03	8012	/SIDS2/cgcgdata/geneseq/geneseq/NA1996.DAT:AAZ37329	-	36.00	77.42	3.1e+04	2880
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ57309	-	37.00	93.29	4.0e+03	8418	/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ92474	-	36.00	77.42	3.1e+04	2880
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ57308	-	37.00	92.85	4.3e+03	8798	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ98812	-	36.00	77.42	3.1e+04	2880
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH67704	-	37.00	92.65	4.4e+03	8979	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ98147	-	36.00	77.42	3.1e+04	2880
/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ90812	-	37.00	92.58	4.6e+03	9048	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH13530	-	36.00	71.68	6.4e+04	5147
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ08923	-	37.00	90.78	5.6e+03	10847	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ16160	-	36.00	71.68	6.4e+04	5147
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ08924	-	37.00	90.73	5.6e+03	10900	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ22303	-	36.00	52.71	1.0e+05	3499
/SIDS2/cgcgdata/geneseq/geneseq/NA1997.DAT:AAZ74322	-	37.00	90.40	5.9e+03	11271	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ81490	-	36.00	47.19	1.0e+06	6115
/SIDS2/cgcgdata/geneseq/geneseq/NA1994.DAT:AAZ67902	-	37.00	81.87	1.7e+04	26700	/SIDS2/cgcgdata/geneseq/geneseq/NA1996.DAT:AAZ32019	-	35.50	118.34	158.39	374
/SIDS2/cgcgdata/geneseq/geneseq/NA1998.DAT:AAZ15946	-	37.00	81.87	1.7e+04	26700	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ45242	-	35.50	118.54	158.39	374
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ81283	-	37.00	81.87	1.7e+04	26700	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAH13530	-	35.50	114.08	280.63	587
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ81453	-	37.00	78.79	2.6e+04	36471	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ39349	-	35.50	106.64	729.29	1246
/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ06831	-	37.00	78.39	2.7e+04	37948	/SIDS2/cgcgdata/geneseq/geneseq/NA1995.DAT:AAZ25098	-	35.00	129.36	39.59	104
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ55887	-	37.00	72.51	5.7e+04	68750	/SIDS2/cgcgdata/geneseq/geneseq/NA2001.DAT:AAZ3208	-	35.00	119.42	141.63	284
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ29349	-	37.00	72.06	6.1e+04	71989	/SIDS2/cgcgdata/geneseq/geneseq/NA1999.DAT:AAZ12870	-	35.00	118.87	151.83	300
/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ21611	-	37.00	56.41	4.1e+05	349980	/SIDS2/cgcgdata/geneseq/geneseq/NA2000.DAT:AAZ74488	-	35.00	117.84	173.32	333



/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI09561 +	34.00	125.01	69.13	111	1	/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAQ05791 +	34.00	97.63	2.3e+03	1767
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ4744 -	34.00	120.97	116.08	167	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAS08953 +	34.00	97.63	2.3e+03	1767
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAF12741 -	34.00	117.34	184.88	241	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAS1470 -	34.00	97.53	2.3e+03	1785
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAF17842 -	34.00	117.34	183.88	241	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC4198 -	34.00	97.00	2.5e+03	1884
/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH89912 +	34.00	115.51	234.81	290	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC7125 +	34.00	96.94	2.5e+03	1894
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH90749 +	34.00	115.37	237.91	294	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC38724 +	34.00	96.88	2.5e+03	1906
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH90862 +	34.00	114.60	262.82	318	1	/SIDS2/cgdata/geneseq/geneseq/NA1994.DAT:AAV12198 -	34.00	96.83	2.6e+03	1915
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC3679 -	34.00	114.50	265.97	321	1	/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAAT48241 +	34.00	96.81	2.6e+03	1920
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC2921 +	34.00	113.93	286.10	340	1	/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAQ05793 +	34.00	96.81	2.6e+03	1920
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC6520 -	34.00	113.59	298.98	352	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC08955 +	34.00	96.81	2.6e+03	1920
/SIDS2/cgdata/geneseq/geneseq/NA1993.DAT:AAQ61331 -	34.00	113.07	319.60	371	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC82674 -	34.00	96.80	2.6e+03	1922
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ94108 +	34.00	112.33	351.62	400	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI60910 +	34.00	96.75	2.6e+03	1932
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI14780 +	34.00	112.06	363.94	411	1	/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAAT71312 +	34.00	96.72	2.6e+03	1938
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI36138 +	34.00	112.06	363.94	411	1	/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAAF48127 -	34.00	96.72	2.6e+03	1938
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI04573 +	34.00	112.06	363.94	411	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT21928 -	34.00	96.72	2.6e+03	1938
/SIDS2/cgdata/geneseq/geneseq/NA1994.DAT:AAQ79277 +	34.00	111.34	399.12	442	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAT77154 -	34.00	96.72	2.6e+03	1938
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI11584 +	34.00	110.73	431.47	470	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAH59084 +	34.00	96.59	2.6e+03	1962
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI32875 +	34.00	110.73	431.47	470	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI59124 +	34.00	96.55	2.7e+03	1970
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI01508 +	34.00	110.73	431.47	470	1	/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAAT84010 +	34.00	96.19	2.8e+03	2042
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT77002 +	34.00	110.69	433.80	472	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC89915 +	34.00	96.10	2.8e+03	2064
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT59770 +	34.00	108.75	556.03	574	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF33255 +	34.00	95.98	2.9e+03	2087
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT75859 +	34.00	110.50	444.32	481	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT95797 -	34.00	95.98	2.9e+03	2087
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT07671 -	34.00	110.46	446.66	483	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT95797 -	34.00	95.69	3.0e+03	2150
/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH89557 +	34.00	110.12	466.71	500	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF32732 +	34.00	95.69	3.0e+03	2150
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ80497 +	34.00	108.20	596.90	607	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF57479 +	34.00	95.47	3.1e+03	2198
/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAV9621 +	34.00	108.15	606.65	610	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAAF4512 +	34.00	95.47	3.1e+03	2198
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ96425 +	34.00	106.75	719.14	703	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC9383 +	34.00	95.36	3.1e+03	2222
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ21566 +	34.00	107.89	546.22	566	1	/SIDS2/cgdata/geneseq/geneseq/NA1991.DAT:AAH14235 +	34.00	94.45	3.5e+03	2436
/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ80147 +	34.00	107.69	637.11	639	1	/SIDS2/cgdata/geneseq/geneseq/NA1995.DAT:AAQ97381 -	34.00	94.45	3.5e+03	2437
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI17279 +	34.00	107.69	637.11	639	1	/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAAT99802 -	34.00	94.45	3.5e+03	2437
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI57970 +	34.00	107.66	639.64	641	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH88360 -	34.00	94.45	3.5e+03	2437
/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH89484 +	34.00	107.28	671.46	666	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH88361 -	34.00	94.45	3.5e+03	2437
/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAZ96425 +	34.00	106.75	719.14	703	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH88368 -	34.00	94.45	3.5e+03	2437
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ80147 +	34.00	106.43	750.43	727	1	/SIDS2/cgdata/geneseq/geneseq/NA1991.DAT:AAH14235 +	34.00	94.39	3.5e+03	2451
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC30830 +	34.00	106.42	750.43	727	1	/SIDS2/cgdata/geneseq/geneseq/NA1996.DAT:AAAT04426 +	34.00	94.35	3.5e+03	2461
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC40930 +	34.00	106.23	768.82	741	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT69566 +	34.00	94.35	3.5e+03	2461
/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAAT98680 -	34.00	105.85	805.86	769	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT69562 -	34.00	94.28	3.6e+03	2480
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH06941 +	34.00	105.06	893.26	834	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAT69562 -	34.00	94.28	3.6e+03	2480
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH48498 +	34.00	104.70	935.59	865	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH14637 +	34.00	93.84	3.8e+03	2591
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC33243 +	34.00	104.70	935.59	865	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH14637 +	34.00	93.84	3.8e+03	2591
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC1927 -	34.00	104.56	952.09	877	1	/SIDS2/cgdata/geneseq/geneseq/NA1996.DAT:AAAT33487 +	34.00	93.84	3.8e+03	2593
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC98760 +	34.00	103.73	1.1e+03	954	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC89923 +	34.00	93.46	4.0e+03	2692
/SIDS2/cgdata/geneseq/geneseq/NA1993.DAT:AAAC3159 -	34.00	103.01	1.2e+03	1026	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC89922 +	34.00	93.12	4.1e+03	2788
/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ34526 -	34.00	102.97	1.2e+03	1030	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH54984 +	34.00	93.06	4.2e+03	2803
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH84649 +	34.00	102.69	1.2e+03	1059	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF93884 +	34.00	92.82	4.3e+03	2872
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC33243 +	34.00	102.33	1.3e+03	1099	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAD03266 -	34.00	92.73	4.3e+03	2900
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC52116 -	34.00	102.21	1.3e+03	1112	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH14658 +	34.00	92.59	4.4e+03	2942
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC53685 +	34.00	101.89	1.3e+03	1149	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH17711 +	34.00	92.58	4.4e+03	2944
/SIDS2/cgdata/geneseq/geneseq/NA1994.DAT:AAQ54832 +	34.00	101.24	1.5e+03	1227	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAH51753 +	34.00	92.39	4.5e+03	3001
/SIDS2/cgdata/geneseq/geneseq/NA1996.DAT:AAAT37306 +	34.00	101.24	1.5e+03	1227	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAH54052 +	34.00	92.39	4.5e+03	3002
/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAAT37393 +	34.00	100.94	1.5e+03	1265	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAZ12184 +	34.00	92.25	4.6e+03	3045
/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAAT98643 -	34.00	100.85	1.5e+03	1276	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ12185 +	34.00	92.25	4.6e+03	3045
/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAAT30741 +	34.00	100.80	1.5e+03	1282	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ34518 -	34.00	92.25	4.6e+03	3045
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAF67128 +	34.00	100.72	1.6e+03	1293	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ34518 -	34.00	91.74	4.9e+03	3205
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF44756 -	34.00	100.72	1.6e+03	1293	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAF81716 +	34.00	91.49	5.1e+03	3287
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC59516 -	34.00	100.54	1.6e+03	1316	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH54214 +	34.00	91.29	5.2e+03	3355
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC59516 -	34.00	100.54	1.6e+03	1316	1	/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAQ05794 +	34.00	90.77	5.6e+03	3534
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC59516 -	34.00	100.54	1.6e+03	1316	1	/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAQ05794 +	34.00	90.77	5.6e+03	3534
/SIDS2/cgdata/geneseq/geneseq/NA1995.DAT:AAAT35799 +	34.00	99.87	1.7e+03	1409	1	/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAQ05795 +	34.00	90.77	5.6e+03	3534
/SIDS2/cgdata/geneseq/geneseq/NA1990.DAT:AAAC02854 -	34.00	99.13	1.9e+03	1519	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAS08957 +	34.00	90.77	5.6e+03	3534
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC02854 -	34.00	99.13	1.9e+03	1520	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAS08957 +	34.00	90.77	5.6e+03	3534
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC50759 -	34.00	99.08	1.9e+03	1526	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI60356 -	34.00	90.70	5.6e+03	3557
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC36427 -	34.00	99.07	1.9e+03	1528	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI58570 -	34.00	90.65	5.7e+03	3577
/SIDS2/cgdata/geneseq/geneseq/NA1991.DAT:AAQ11798 +	34.00	98.88	2.0e+03	1557	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH19184 +	34.00	90.29	5.9e+03	3712
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC90013 +	34.00	98.84	2.0e+03	1563	1	/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAZ57339 +	34.00	90.00	6.2e+03	3820
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAC90013 +	34.00	98.84	2.0e+03	1563	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ57339 +	34.00	89.74	6.4e+03	3924
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC66440 +	34.00	98.74	2.0e+03	1579	1	/SIDS2/cgdata/geneseq/geneseq/NA1999.DAT:AAZ57339 +	34.00	89.73	6.4e+03	3927
/SIDS2/cgdata/geneseq/geneseq/NA1991.DAT:AAQ11799 +	34.00	98.18	2.2e+03	1671	1	/SIDS2/cgdata/geneseq/geneseq/NA1998.DAT:AAV54955 +	34.00	87.85	8.1e+03	4746
/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAH353607 +	34.00	98.02	2.2e+03	1698	1	/SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAAT03901 +	34.00	87.85	8.1e+03	4746
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC37576 +	34.00	97.67	2.3e+03	1759	1	/SIDS2/cgdata/geneseq/geneseq/NA1996.DAT:AAAT03901 +	34.00	87.85	8.1e+03	4747
/SIDS2/cgdata/geneseq/geneseq/NA2000.DAT:AAAC37576 +	34.00	97.66	2.3e+03	1762	1	/SIDS2/cgdata/geneseq/geneseq/NA1997.DAT:AAV74417 +	34.00	87.63	8.3e+03	4854









/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF29701	-	33.00	93.32	4.0e+03	1879	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH02905	-	33.00	86.72	9.4e+03	3663
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH98391	-	33.00	93.26	4.1e+03	1891	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX9689	+	33.00	86.71	9.4e+03	3664
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH9696	-	33.00	93.26	4.1e+03	1891	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX77013	+	33.00	86.71	9.4e+03	3666
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH9699	-	33.00	93.26	4.1e+03	1891	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAAX77012	+	33.00	86.70	9.4e+03	3667
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH8788	-	33.00	93.24	4.1e+03	1894	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV59424	+	33.00	86.39	9.8e+03	3786
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH15644	+	33.00	93.23	4.1e+03	1897	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ89988	+	33.00	86.39	9.8e+03	3786
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH93817	+	33.00	93.22	4.1e+03	1898	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AACT7795	+	33.00	86.39	9.8e+03	3787
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH20309	+	33.00	93.20	4.1e+03	1902	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX20634	+	33.00	86.21	1.0e+04	3856
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH22396	-	33.00	93.15	4.1e+03	1911	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAC4933	+	33.00	86.00	1.0e+04	3936
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH297138	-	33.00	93.14	4.1e+03	1913	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX87639	+	33.00	85.87	1.0e+04	3988
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH41359	-	33.00	93.12	4.1e+03	1917	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF72857	-	33.00	85.87	1.0e+04	3988
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH97093	-	33.00	93.05	4.2e+03	1931	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI158250	-	33.00	85.55	1.1e+04	4139
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17106	+	33.00	92.87	4.3e+03	1966	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI50464	+	33.00	85.51	1.1e+04	4150
/SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:AAH91045	+	33.00	92.84	4.3e+03	1973	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAD07790	-	33.00	85.03	1.2e+04	4344
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH55595	+	33.00	92.80	4.3e+03	1980	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX39935	+	33.00	84.79	1.2e+04	4451
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH55024	+	33.00	92.67	4.4e+03	2007	/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAAT35186	-	33.00	84.74	1.2e+04	4473
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH45226	+	33.00	92.26	4.6e+03	2091	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAH32152	+	33.00	84.00	1.3e+04	4820
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH47514	+	33.00	92.18	4.7e+03	2109	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH62756	+	33.00	84.00	1.3e+04	4821
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH73761	+	33.00	92.17	4.7e+03	2111	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH57430	-	33.00	83.63	1.4e+04	5005
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH15851	+	33.00	92.13	4.7e+03	2118	/SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:AAQ27764	+	33.00	83.46	1.4e+04	5089
/SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:AAO06594	+	33.00	92.07	4.7e+03	2132	/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV74643	+	33.00	83.37	1.4e+04	5136
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH232138	+	33.00	91.97	4.8e+03	2153	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAH8984	+	33.00	83.14	1.5e+04	5258
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH89575	+	33.00	91.97	4.8e+03	2153	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH38280	-	33.00	83.13	1.5e+04	5264
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH289587	+	33.00	91.97	4.8e+03	2153	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH28277	-	33.00	83.03	1.5e+04	5317
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH31531	+	33.00	91.82	4.9e+03	2187	/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAAT51085	+	33.00	83.03	1.5e+04	5319
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH87381	+	33.00	91.48	5.1e+03	2263	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF76289	+	33.00	82.99	1.5e+04	5340
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAH35307	+	33.00	91.42	5.1e+03	2276	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF72846	+	33.00	82.79	1.6e+04	5446
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH94471	-	33.00	91.39	5.2e+03	2284	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAH4049	+	33.00	82.38	1.6e+04	5680
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH76389	-	33.00	91.26	5.2e+03	2314	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV84328	+	33.00	82.02	1.7e+04	5889
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH76390	-	33.00	91.13	5.3e+03	2344	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ43720	+	33.00	81.77	1.8e+04	6038
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH76635	+	33.00	91.12	5.3e+03	2347	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAH44592	+	33.00	81.50	1.8e+04	6204
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH00037	+	33.00	90.99	5.4e+03	2379	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH16666	-	33.00	81.34	1.9e+04	6304
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH76577	+	33.00	90.97	5.4e+03	2382	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH60792	-	33.00	80.87	2.0e+04	6616
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH64360	+	33.00	90.97	5.4e+03	2382	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV19051	-	33.00	80.86	2.0e+04	6623
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17380	+	33.00	90.78	5.6e+03	2430	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV19053	-	33.00	80.86	2.0e+04	6623
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH34932	-	33.00	90.73	5.6e+03	2440	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX4013	-	33.00	80.12	2.2e+04	7131
/SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:AAO03665	+	33.00	90.63	5.7e+03	2465	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV34061	-	33.00	79.81	2.3e+04	7358
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI62735	-	33.00	90.47	5.8e+03	2506	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI59158	+	33.00	79.67	2.3e+04	7518
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17822	-	33.00	90.32	5.9e+03	2544	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ43718	+	33.00	79.47	2.4e+04	7622
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI51247	+	33.00	90.30	5.9e+03	2550	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ27543	+	33.00	79.45	2.4e+04	7633
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAO33136	+	33.00	90.11	6.1e+03	2600	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAC68230	+	33.00	79.45	2.4e+04	7633
/SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:AAO26664	+	33.00	89.96	6.2e+03	2640	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX98537	+	33.00	79.45	2.4e+04	7633
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ49227	+	33.00	89.93	6.2e+03	2648	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAO30370	-	33.00	79.39	2.4e+04	7680
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAH61539	-	33.00	89.53	6.5e+03	2755	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH76578	+	33.00	79.13	2.5e+04	7885
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ22704	+	33.00	89.47	6.6e+03	2773	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AACT5460	+	33.00	78.95	2.5e+04	8028
/SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAH45353	+	33.00	89.41	6.6e+03	2790	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX3523	+	33.00	78.49	2.7e+04	8416
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAH97303	+	33.00	89.11	6.9e+03	2876	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAT28521	+	33.00	77.71	3.0e+04	9100
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH02917	+	33.00	89.11	6.9e+03	2876	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV52192	+	33.00	77.58	3.0e+04	9223
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAH18573	+	33.00	88.88	7.1e+03	2944	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX29064	+	33.00	77.50	3.0e+04	9301
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAH05112	-	33.00	88.59	7.4e+03	3031	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ36256	-	33.00	76.41	3.5e+04	1038
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAH060305	-	33.00	88.59	7.4e+03	3031	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI12991	+	33.00	75.60	3.9e+04	1126
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH54296	+	33.00	88.53	7.4e+03	3048	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX20649	-	33.00	73.65	4.9e+04	1372
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH64361	-	33.00	88.37	7.6e+03	3099	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAI13166	-	33.00	73.59	5.0e+04	1381
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH76565	+	33.00	88.28	7.7e+03	3127	/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAAT1418	+	33.00	71.48	6.5e+04	1708
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAH02847	+	33.00	88.23	7.7e+03	3142	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX23524	-	33.00	70.79	7.1e+04	1831
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH28278	+	33.00	88.23	7.7e+03	3144	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAAX23560	-	33.00	70.10	7.7e+04	1963
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH24331	-	33.00	88.21	7.8e+03	3149	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV77439	+	33.00	69.38	8.5e+04	2112
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17293	-	33.00	88.17	7.8e+03	3161	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX29063	-	33.00	64.36	1.6e+05	3510
/SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAO66685	+	33.00	88.09	7.9e+03	3188	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAX29064	+	33.00	64.15	1.6e+05	3582
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH18573	+	33.00	88.07	7.9e+03	3193	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ45596	+	33.00	60.93	2.4e+05	4964
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH72593	+	33.00	87.91	8.1e+03	3246	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF28545	-	33.00	60.00	2.7e+05	5454
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH39764	+	33.00	87.76	8.2e+03	3296	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAH8954	+	33.00	58.59	3.2e+05	6290
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH54461	+	33.00	87.74	8.2e+03	3302	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF28505	+	33.00	57.93	3.4e+05	6721
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH159461	+	33.00	87.67	8.3e+03	3327	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH08693	-	33.00	54.04	5.3e+05	9962
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAH39793	-	33.00	87.61	8.4e+03	3345	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV19941	-	33.00	53.10	5.8e+05	1095
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17490	-	33.00	87.58	8.4e+03	3355	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH62689	-	33.00	50.85	7.3e+05	1375
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH48466	+	33.00	87.52	8.5e+03	3378	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAF21607	-	33.00	42.96	1.4e+06	3051
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH55029	+	33.00	87.46	8.5e+03	3398	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAAF1607	-	33.00	42.83	1.4e+06	3094
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH21754	+	33.00	87.44	8.6e+03	3405	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH61508	-	33.00	41.61	1.5e+06	3499
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAH16236	+	33.00	87.12	8.6e+03	3515	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAAF28525	-	33.00	41.61	1.5e+06	3499
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH34535	+	33.00	87.10	8.9e+03	3522	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH68528	+	33.00	41.61	1.5e+06	3499
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH54572	+	33.00	86.86	9.2e+03	3609	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH68530	+	33.00	41.61	1.5e+06	3499

```
/SID22/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH68530 - 33.00 41.61 1.5e+06 349980
/SID22/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI61373 - 33.00 37.81 1.7e+06 513445
/SID22/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV21209 - 33.00 26.18 1.9e+06 1664976
/SID22/gcgdata/geneseq/geneseqn/NA1996.DAT:AAH42063 + 33.00 25.24 1.9e+06 1830121
/SID22/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH42830 + 32.50 108.27 591.93 344
/SID22/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH37057 + 32.50 107.05 691.86 389
/SID22/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV77529 + 32.50 104.83 920.09 487
/SID22/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH10907 + 32.50 103.01 1.2e+03 585
/SID22/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH82204 + 32.50 102.71 1.2e+03 603
/SID22/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH32864 + 32.50 99.83 1.7e+03 807
/SID22/gcgdata/geneseq/geneseqn/NA2000.DAT:AAH45203 + 32.50 99.30 1.9e+03 851
```

seq\_name: /SID22/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH36284

seq\_documentation\_block:  
ID AAX36284 standard; DNA; 51 BP.

XX  
AC AAX36284;  
DT 16-JUL-1999 (first entry)  
DE Oligonucleotide derived from the wheat rht gene.  
KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.

OS Triticum aestivum.  
XX  
PN WO9909174-A1.  
XX  
PD 25-FEB-1999.  
XX  
PF 07-AUG-1998; 9BWO-GB02383.  
XX  
PR 13-AUG-1997; 97GB-0017192.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Harberd NP, Peng J, Richards DE;  
XX  
DR WPI; 1999-181040/15.  
XX

PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype

XX  
PS Claim 3; Page 53; 88pp; English.  
XX  
CC The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence is  
CC derived from the wheat rht gene.

XX  
SQ Sequence 51 BP; 7 A; 16 C; 21 G; 7 T; 0 other;

alignment\_scores:  
Quality: 81.00 Length: 17  
Ratio: 4.765 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-485-529-104 x AAX36284 ..  
Align seg 1/1 to: AAX36284 from: 1 to: 51  
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
||||| 1 GAGGAGCTGCTGGCGCGCTCGGGTACAAGGTGCGCGCTCCGACATGCC 50  
17 a 17  
51 G 51

seq\_name: /SID22/gcgdata/geneseq/geneseqn/NA1999.DAT:AAH36263

seq\_documentation\_block:  
ID AAX36263 standard; DNA; 453 BP.

XX  
AC AAX36263;  
DT 16-JUL-1999 (first entry)  
DE DNA sequence obtained after sequencing wheat Rht clone 14a1.  
XX  
KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.  
XX  
OS Triticum aestivum.  
XX  
PN WO9909174-A1.  
XX  
PD 25-FEB-1999.  
XX  
PF 07-AUG-1998; 9BWO-GB02383.  
XX  
PR 13-AUG-1997; 97GB-0017192.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Harberd NP, Peng J, Richards DE;  
XX  
DR WPI; 1999-181040/15.  
XX

PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype

XX  
PS Disclosure; Fig 2b(9); 88pp; English.  
XX  
CC The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence was  
CC obtained after partially sequencing wheat Rht clone 14a1.

XX  
SQ Sequence 453 BP; 85 A; 136 C; 158 G; 67 T; 7 other;

alignment\_scores:  
Quality: 81.00 Length: 17  
Ratio: 4.765 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-485-529-104 x AAX36263 ..  
 Align seg 1/1 to: AAX36263 from: 1 to: 453  
 1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||  
 273 GACGAGCTGCTGGCGGCTGGGTACAGGTGGCGCTCCGACATGGC 322  
 17 a 17  
 |  
 323 G 323

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36279

seq\_documentation\_block:  
 ID AAX36279 standard; DNA; 2125 BP.  
 XX  
 AC AAX36279;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE Wheat Rht clone 5a1 genomic sequence.  
 XX  
 KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 PN WO9909174-A1.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 07-AUG-1998; 98WO-GB02383.  
 XX  
 PR 13-AUG-1997; 97GB-0017192.  
 XX  
 PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX  
 PI Harberd NP, Peng J, Richards DE;  
 XX  
 DR WPI; 1999-181040/15.  
 DR P-PSDB; AAY02540.  
 XX  
 XX New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX  
 PS Disclosure; Fig 8a; 88pp; English.  
 XX  
 CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the wheat Rht clone 5a1 genomic sequence.  
 XX  
 SQ Sequence 2125 BP; 325 A; 768 C; 723 G; 309 T; 0 other;

alignment\_scores:  
 Quality: 81.00 Length: 17  
 Ratio: 4.765 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-485-529-104 x AAX36279 ..  
 Align seg 1/1 to: AAX36279 from: 1 to: 2125  
 1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||  
 139 GACGAGCTGCTGGCGGCTGGGTACAGGTGGCGCTCCGACATGGC 188  
 17 a 17  
 |  
 189 G 189

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36275

seq\_documentation\_block:  
 ID AAX36275 standard; DNA; 2709 BP.  
 XX  
 AC AAX36275;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE Composite DNA sequence of wheat Rht gene.  
 XX  
 KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 PN WO9909174-A1.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 07-AUG-1998; 98WO-GB02383.  
 XX  
 PR 13-AUG-1997; 97GB-0017192.  
 XX  
 PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX  
 PI Harberd NP, Peng J, Richards DE;  
 XX  
 DR WPI; 1999-181040/15.  
 XX  
 XX New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX  
 PS Disclosure; Fig 3a; 88pp; English.  
 XX  
 CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the composite DNA sequence of wheat Rht gene.  
 XX  
 SQ Sequence 2709 BP; 473 A; 900 C; 823 G; 482 T; 31 other;

alignment\_scores:  
 Quality: 81.00 Length: 17  
 Ratio: 4.765 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

```
alignment_block:
US-09-485-529-104 x AAX36275  ..
Align seg 1/1 to: AAX36275 from: 1 to: 2709

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
273 GAGGAGCTGCTGGCGGCTCGGGTACAGGTGCGGCTCCGACATGGC 322

17 a 17
|
323 G 323

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36281

seq_documentation_block:
ID AAX36281 standard; DNA; 302 BP.
XX
AC AAX36281;
XX
DT 16-JUL-1999 (first entry)
XX
DE Partial sequence of the maize D8-1 allele.
XX
KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;
KW paclobutrazol; maize; ss.
XX
OS Zea mays.
XX
PN WO9909174-Al.
XX
PD 25-FEB-1999.
XX
PF 07-AUG-1998; 98WO-GB02383.
XX
PR 13-AUG-1997; 97GB-0017192.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Harberd NP, Peng J, Richards DE;
XX
DR WPI; 1999-181040/15.
DR P-PSDB; AAY02542.
XX
New Triticum Aestivum polynucleotides - encode a polypeptide which
PT provides inhibition of the growth of plants, which inhibition is
PT antagonised by gibberellin, used to confer a dwarf phenotype
XX
PS Disclosure; Fig 11a; 88pp; English.
XX
CC The specification describes polypeptides encoded by the Rht gene (and
CC its homologues) that, when expressed in Triticum Aestivum, inhibit
CC growth of the plant. This growth inhibition is antagonised by
CC gibberellin. The products can be used to provide Rht expression in
CC plants, conferring a dwarf phenotype on a plant which is correctable
CC by treatment with gibberellin. In addition, the products can be
CC used to produce Rht mutant plants which are dwarfed compared with
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants
CC may be made by knocking out Rht or the relevant homologous gene in
CC the plant of interest. Plants may be made which are resistant to
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds
CC dwarf but let crop plants grow tall. The present sequence represents
CC the partial sequence of the maize D8-1 allele.
XX
SQ Sequence 302 BP; 52 A; 82 C; 123 G; 45 T; 0 other;
```

```
alignment_scores:
Quality: 78.00 Length: 17
Ratio: 4.588 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118
```

```
alignment_block:
US-09-485-529-104 x AAX36281  ..
Align seg 1/1 to: AAX36281 from: 1 to: 302

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
100 GATGAGCTGCTGGCGGCTCGGGTACAGGTGCGTTCGATGCGATGGC 149

17 a 17
|
150 G 150

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36282

seq_documentation_block:
ID AAX36282 standard; DNA; 371 BP.
XX
AC AAX36282;
XX
DT 16-JUL-1999 (first entry)
XX
DE Partial sequence of the maize D8-2023 allele.
XX
KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;
KW paclobutrazol; maize; ss.
XX
OS Zea mays.
XX
PN WO9909174-Al.
XX
PD 25-FEB-1999.
XX
PF 07-AUG-1998; 98WO-GB02383.
XX
PR 13-AUG-1997; 97GB-0017192.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Harberd NP, Peng J, Richards DE;
XX
DR WPI; 1999-181040/15.
DR P-PSDB; AAY02543.
XX
New Triticum Aestivum polynucleotides - encode a polypeptide which
PT provides inhibition of the growth of plants, which inhibition is
PT antagonised by gibberellin, used to confer a dwarf phenotype
XX
PS Disclosure; Fig 11c; 88pp; English.
XX
CC The specification describes polypeptides encoded by the Rht gene (and
CC its homologues) that, when expressed in Triticum Aestivum, inhibit
CC growth of the plant. This growth inhibition is antagonised by
CC gibberellin. The products can be used to provide Rht expression in
CC plants, conferring a dwarf phenotype on a plant which is correctable
CC by treatment with gibberellin. In addition, the products can be
CC used to produce Rht mutant plants which are dwarfed compared with
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants
CC may be made by knocking out Rht or the relevant homologous gene in
CC the plant of interest. Plants may be made which are resistant to
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds
CC dwarf but let crop plants grow tall. The present sequence represents
CC the partial sequence of the maize D8-2023 allele.
XX
SQ Sequence 371 BP; 56 A; 121 C; 139 G; 55 T; 0 other;
```

```
alignment_scores:
Quality: 78.00 Length: 17
Ratio: 4.588 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x AAX36282 ..

Align seg 1/1 to: AAX36282 from: 1 to: 371

1 AspCluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||  
 64 GATGAGCTGCTGCGCGCTCGGGTACAGGTGCTGCGGATATGGC 113  
 17 a 17  
 ||  
 114 G 114

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36277

seq\_documentation\_block:  
 ID AAX36277 standard; DNA; 770 BP.

XX AC AAX36277;

XX DT 16-JUL-1999 (first entry)

XX DE Rice EST D39460 sequence, homologous to wheat Rht gene.

XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 XX KW paclobutrazol; rice; expressed sequence tag; EST; ss.

XX OS Oryza sativa.

XX PN WO9909174-Al.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-GB02383.

XX PR 13-AUG-1997; 97GB-0017192.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Harberd NP, Peng J, Richards DE;

XX DR WPI; 1999-181040/15.

XX DR P-PSDB; AAY02538.

XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
 provides inhibition of the growth of plants, which inhibition is  
 antagonised by gibberellin, used to confer a dwarf phenotype

XX PS Claim 13; Fig 6a; 88pp; English.

XX CC The specification describes polypeptides encoded by the Rht gene (and  
 its homologues) that, when expressed in Triticum Aestivum, inhibit  
 growth of the plant. This growth inhibition is antagonised by  
 gibberellin. The products can be used to provide Rht expression in  
 plants, conferring a dwarf phenotype on a plant which is correctable  
 by treatment with gibberellin. In addition, the products can be  
 used to produce Rht mutant plants which are dwarfed compared with  
 wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 may be made by knocking out Rht or the relevant homologous gene in  
 the plant of interest. Plants may be made which are resistant to  
 compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 dwarf but let crop plants grow tall. The present sequence represents  
 rice expressed sequence tag (EST) AAB39460, which is homologous to the  
 wheat Rht gene.

XX SQ Sequence 770 BP; 104 A; 252 C; 294 G; 120 T; 0 other;

alignment\_scores:

Quality: 78.00 Length: 17

Ratio: 4.588 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x AAX36277 ..

Align seg 1/1 to: AAX36277 from: 1 to: 770

1 AspCluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||  
 114 GACGAGCTGCTGCGCGCTCGGGTACAGGTGCTGCGCATGGC 163  
 17 a 17  
 ||  
 164 C 164

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36280

seq\_documentation\_block:

ID AAX36280 standard; DNA; 2255 BP.

XX AC AAX36280;

XX DT 16-JUL-1999 (first entry)

XX DE Maize lal genomic clone sequence.

XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 XX KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 XX KW paclobutrazol; maize; ss.

XX OS Zea mays.

XX PN WO9909174-Al.

XX PD 25-FEB-1999.

XX PF 07-AUG-1998; 98WO-GB02383.

XX PR 13-AUG-1997; 97GB-0017192.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Harberd NP, Peng J, Richards DE;

XX DR WPI; 1999-181040/15.

XX DR P-PSDB; AAY02541.

XX PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
 provides inhibition of the growth of plants, which inhibition is  
 antagonised by gibberellin, used to confer a dwarf phenotype

XX PS Disclosure; Fig 9a; 88pp; English.

XX CC The specification describes polypeptides encoded by the Rht gene (and  
 its homologues) that, when expressed in Triticum Aestivum, inhibit  
 growth of the plant. This growth inhibition is antagonised by  
 gibberellin. The products can be used to provide Rht expression in  
 plants, conferring a dwarf phenotype on a plant which is correctable  
 by treatment with gibberellin. In addition, the products can be  
 used to produce Rht mutant plants which are dwarfed compared with  
 wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 may be made by knocking out Rht or the relevant homologous gene in  
 the plant of interest. Plants may be made which are resistant to  
 compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 dwarf but let crop plants grow tall. The present sequence represents  
 the maize lal genomic clone sequence.

XX SQ Sequence 2255 BP; 334 A; 817 C; 737 G; 367 T; 0 other;

alignment\_scores: Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:  
 US-09-485-529-104 x AAX36280 ..

Align seg 1/1 to: AAX36280 from: 1 to: 2255

1 AspGluLeuAlaAlaLeuClTyTyrLysValArgAlaSerAspMetAla 17  
 |||||  
 308 GATGAGCTGCTGCGCGCTCGGGTACAAGGTGCGTTCGTCGATATGCC 357  
 17 a 17  
 358 G 358

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: AAX36276

seq\_documentation\_block:  
 ID AAX36276 standard; DNA; 725 BP.

XX AC AAX36276;  
 XX DT 16-JUL-1999 (first entry)  
 XX DE Partial cDNA sequence of rice D39460 (a wheat Rht gene homologue).  
 XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX OS Oryza sativa.  
 XX PN WO9909174-A1.  
 XX PD 25-FEB-1999.  
 XX PF 07-AUG-1998; 98WO-GB02383.  
 XX PR 13-AUG-1997; 97GB-0017192.  
 XX PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX PI Harberd NP, Peng J, Richards DE;  
 XX WPI; 1999-181040/15.  
 XX PS New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Disclosure; Fig 4a; 88pp; English.  
 XX CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the partial cDNA sequence of rice AAD39460 (a wheat Rht gene homologue).  
 XX SQ Sequence 725 BP; 98 A; 226 C; 276 G; 119 T; 6 other;

alignment\_scores: Quality: 72.00 Length: 16  
 Ratio: 4.500 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 93.750

alignment\_block:  
 US-09-485-529-104 x AAX36276 ..

Align seg 1/1 to: AAX36276 from: 1 to: 725

2 GluLeuLeuAlaAlaLeuClTyTyrLysValArgAlaSerAspMetAla 17  
 |||||  
 109 GAGCTGCTGCGCGCTCGGGTACAAGGTGCGTTCGTCGACATGCC 156

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: AAX36273

seq\_documentation\_block:  
 ID AAX36273 standard; DNA; 324 BP.

XX AC AAX36273;  
 XX DT 16-JUL-1999 (first entry)  
 XX DE DNA sequence obtained after sequencing wheat Rht clone 5a1.  
 XX KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX OS Triticum aestivum.  
 XX PN WO9909174-A1.  
 XX PD 25-FEB-1999.  
 XX PF 07-AUG-1998; 98WO-GB02383.  
 XX PR 13-AUG-1997; 97GB-0017192.  
 XX PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX PI Harberd NP, Peng J, Richards DE;  
 XX WPI; 1999-181040/15.  
 XX PS New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX Disclosure; Fig 2c(6); 88pp; English.  
 XX CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence was  
 CC obtained after partially sequencing wheat Rht clone 5a1.  
 XX SQ Sequence 324 BP; 39 A; 141 C; 85 G; 56 T; 3 other;

alignment\_scores: Quality: 70.00 Length: 15  
 Ratio: 4.667 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000



```
alignment_block:
US-09-485-529-104 x AAX36273/rev ..
Align seg 1/1 to reverse of: AAX36273 from: 1 to: 324
3 LeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
|||||
156 CTGCTGGCGGCGCTCGGTACAAAGTGGCGCTCCGACATGGCG 112

seq_name: /SDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC65312
seq_documentation_block:
ID AAC65312 standard; cDNA; 457 BP.
XX
AC AAC65312;
XX
DT 09-FEB-2001 (first entry)
XX
DE Arabidopsis SCL EST T22782.
XX
KW Scarecrow gene; SCR gene; plant; regulatory gene; SCARECROW-like; SCL;
KW transgenic plant; cell division; molecular marker; herbicide resistance;
KW salt resistance; pathogen resistance; insect resistance; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200053723-A2.
XX
PD 14-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-US05875.
XX
PR 10-MAR-1999; 99US-0265585.
XX
PA (UUNY ) UNIV NEW YORK STATE.
XX
PI Benfey PN, Di Lorenzo L, Wysocka-Diller J, Malamy JE, Pysh L;
PI Helariutta Y, Bruce W, Lim J;
XX
DR WPI; 2000-594315/56.
XX
PT Scarecrow gene useful for producing transgenic plants expressing genes
PT whose product increases starch, lignin or cellulose biosynthesis and
PT confers herbicide, pathogen or insect resistance .
XX
PS Claim 6; Fig 28M; 200pp; English.
XX
CC The present sequence is given in a specification relating to the
CC structure and function of a regulatory gene, SCARECROW (SCR).
CC SCARECROW-like (SCL) genes encoding proteins containing an amino acid
CC sequence similar to the sequence of MOTIF III (VHIIID) of Arabidopsis SCR
CC protein are disclosed. SCR and SCL genes are useful for producing
CC transgenic plants whose cell division is modified and root and/or stem
CC development and gravitropism of stem or hypocotyl is altered. Cell
CC division is increased or decreased in roots resulting in thicker or
CC thinner root development. The transgenic plants are useful for expressing
CC a gene of interest encoding a gene product that confers herbicide, salt,
CC pathogen or insect resistance in root or embryos and genes encoding
CC starch, lignin or cellulose biosynthesis in shoots. The SCR gene also
CC confers less susceptibility to lodging in the transgenic plants than a
CC wild-type plant. SCR gene sequences are also useful as molecular markers
CC for a quantitative trait e.g. root or gravitropism trait in molecular
CC breeding of crop plants.
XX
SQ Sequence 457 BP; 141 A; 86 C; 87 G; 125 T; 18 other;
```

```
alignment_scores:
Quality: 70.00 Length: 17
Ratio: 4.118 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353
```

```
alignment_block:
US-09-485-529-104 x AAC65312
Align seg 1/1 to: AAC65312 from: 1 to: 457
1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
|||||
256 GATGAGCTTCTAGCTGTCTTGGTTATAGGTAGGTAGTCCGAAATGCG 305

17 a 17
306 T 306

seq_name: /SDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC45745
seq_documentation_block:
ID AAC45745 standard; DNA; 1602 BP.
XX
AC AAC45745;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47609.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
```



PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146380.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

alignment\_scores:  
Quality: 70.00

Length: 17

Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353  
alignment\_block:  
US-09-485-529-104 x AAC45745 ..  
Align seg 1/1 to: AAC45745 from: 1 to: 1602  
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
82 GATGAGCTTCTAGCTGTTCTGGTTACAAAGTTAGGTTCATCCGAAATGCC 131  
17 a 17  
132 T 132

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAD06661

seq\_documentation\_block:

ID AAD06661 standard; cDNA; 1764 BP.

XX AAD06661;

DT 10-AUG-2001 (first entry)

DE A. thaliana transcription factor G308 homolog, G307 cDNA.

XX Plant transcription factor; phenotype: sugar sensing characteristic;  
KW transgenic plant; plant yield; growth; germination; photosynthesis;  
KW glyoxylate metabolism; respiration; pathogen response; wounding response;  
KW cell cycle regulation; pigmentation; flowering; senescence; physiology;  
KW storage organ; metabolism; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT CDS 1..1764

FT /\*tag= a

FT /product= "Transcription factor homolog"

XX WO200135725-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31414.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (JIAN/) JIANG C.

PA (HEAR/) HEARD J.

PA (PINE/) PINEDA O.

PA (PILG/) PILGRIM M.

PA (ADAM/) ADAM L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (SAMA/) SAMAHA R.

XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;

PI Yu G, Samaha R;

XX WPI; 2001-335977/35.

XX P-PSDB; AAE02560.

XX Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the sugar sensing characteristics of plants and increasing  
PT yield, e.g. corn, potato and cotton plants -

XX Claim 4; Page 113-115; 151pp; English.

XX The patent relates to polynucleotides encoding 35 plant transcription

CC factors which may be used to modify phenotype associated with a plant's  
CC sugar sensing characteristics and increasing yield when their expression  
CC level is altered. Sugars are central regulatory molecules that control  
CC aspects of physiology, metabolism and development. Therefore the cDNAs  
CC and proteins of the invention are useful for modifying the growth and  
CC germination rates of plants, photosynthesis, glyoxylate metabolism,  
CC respiration, starch and sucrose synthesis and degradation, pathogen  
CC response, wounding response, cell cycle regulation, pigmentation,  
CC flowering and senescence of plants and for modifying sink-source  
CC relationships in seeds, tubers, roots, and other storage organs leading  
CC to an increase in yield. The transcription factor polynucleotides and  
CC polypeptides may be used to alter the structure and developmental  
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,  
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,  
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,  
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,  
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,  
CC tobacco, tomato, watermelon, rosaceus fruits and/or vegetable brassicas.  
CC The present sequence is a homolog of Arabidopsis thaliana transcription  
CC factor cDNA.

XX SQ Sequence 1764 BP; 431 A; 383 C; 448 G; 502 T; 0 other;

alignment\_scores:

Quality: 70.00 Length: 17

Ratio: 4.118 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x AAD06661 ..

Align seg 1/1 to: AAD06661 from: 1 to: 1764

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17

|||||

130 GACGAGCTTCTCGCTGTTTACAAAGTTAGGTTCATCCGAGATGCC 179

17 a 17

180 G 180

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAD05776

seq\_documentation\_block:

ID AAD05776 standard; cDNA; 1764 BP.

XX AAD05776;

XX 31-JUL-2001 (first entry)

XX Arabidopsis thaliana transcription factor, G307 cDNA.

XX Transcription factor; biochemical characteristic; controlling element;  
KW structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology; plant trait modification; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT CDS 1..1764

FT /\*tag= a

FT /product= "Transcription factor, G307"

XX WO200136597-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31344.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

```
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (CREE/) CREELMAN R.
PA (YUGG/) YU G.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (HEAR/) HEARD J.
PA (SAMA/) SAMAHA R.
PA (PILG/) PILGRIM M.
PA (PINE/) PINEDA O.
PA (JIAN/) JIANG C.
XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;
PI Pilgrim M, Pineda O, Jiang C;
XX WPI: 2001-335999/35.
DR P-PSDB: AAE01892.
XX Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the biochemical characteristics of plants e.g. corn,
PT potato and cotton plants -
XX Claim 4; Page 66-68; 127pp: English.
XX The present sequence is Arabidopsis thaliana transcription factor,
CC G307 cDNA. The transcription factor is used for altering a plant's
CC biochemical characteristics. The transcription factor may be used to
CC alter the structure and developmental characteristics of plants such as
CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,
CC raspberries, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,
CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,
CC rosaceous fruits and/or vegetable brassicas. Transcription factors are
CC key controlling elements of biological pathways and altering expression
CC levels of 1 or more transcription factors can change entire biological
CC pathways in an organism. Therefore manipulating transcription factor
CC levels in plants offers great potential in agricultural biotechnology
CC for modifying a plant's traits. Transcription factor cDNA is useful in
CC gene therapy.
XX SQ Sequence 1764 BP; 431 A; 383 C; 448 G; 502 T; 0 other;

alignment_scores:
  Quality: 70.00 Length: 17
  Ratio: 4.118 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x AAD05776 ..
Align seg 1/1 to: AAD05776 from: 1 to: 1764
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
130 GACGAGCTTCGCTCTTTAGGTACAAAGTTAGTCTCGGATGCG 179
17 a 17
180 G 180

seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT: AAD06646
seq_documentation_block:
ID_ AAD06646 standard; cDNA; 1951 BP.
XX AAD06646;
XX 10-AUG-2001 (first entry)
XX A. thaliana transcription factor G308 cDNA.
XX
```

```
KW Plant transcription factor; phenotype; sugar sensing characteristic;
KW transgenic plant; plant yield; growth; germination; photosynthesis;
KW glyoxylate metabolism; respiration; pathogen response; wounding response;
KW cell cycle regulation; pigmentation; flowering; senescence; physiology;
KW storage organ; metabolism; ss.
XX Arabidopsis thaliana.
OS
XX Key Location/Qualifiers
FH 196..1794
FT /*tag= a
FT /product= "Transcription factor"
XX WO200135725-A1.
XX 25-MAY-2001.
XX 14-NOV-2000; 2000WO-US1414.
XX 17-NOV-1999; 99US-0166228.
XX 17-APR-2000; 2000US-0197899.
XX 22-AUG-2000; 2000US-0227439.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J.
PA (PINE/) PINEDA O.
PA (PILG/) PILGRIM M.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (SAMA/) SAMAHA R.
XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;
PI Yu G, Samaha R;
XX WPI: 2001-335977/35.
DR P-PSDB: AAE02545.
XX Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the sugar sensing characteristics of plants and increasing
PT yield, e.g. corn, potato and cotton plants -
XX Claim 4; Page 72-74; 151pp: English.
XX The patent relates to polynucleotides encoding 35 plant transcription
CC factors which may be used to modify phenotype associated with a plant's
CC sugar sensing characteristics and increasing yield when their expression
CC level is altered. Sugars are central regulatory molecules that control
CC aspects of physiology, metabolism and development. Therefore the cDNAs
CC and proteins of the invention are useful for modifying the growth and
CC germination rates of plants, photosynthesis, glyoxylate metabolism,
CC respiration, starch and sucrose synthesis and degradation, pathogen
CC response, wounding response, cell cycle regulation, pigmentation,
CC flowering and senescence of plants and for modifying sink-source
CC relationships in seeds, tubers, roots, and other storage organs leading
CC to an increase in yield. The transcription factor polynucleotides and
CC polypeptides may be used to alter the structure and developmental
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.
CC The present sequence is an Arabidopsis thaliana transcription factor
CC cDNA.
XX SQ Sequence 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 other;

alignment_scores:
  Quality: 70.00 Length: 17
  Ratio: 4.118 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 82.353

## alignment\_block:

US-09-485-529-104 x AAD06646 ..

Align seg 1/1 to: AAD06646 from: 1 to: 1951

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
274 GATGAGCTTCTAGCTGTTCTTGGTTACAAAGTTAGGTCATCGGAATGGC 323

17 a 17

1

324 T 324

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAD05791

## seq\_documentation\_block:

ID AAD05791 standard; cDNA; 1951 BP.

AC AAD05791;

DT 31-JUL-2001 (first entry)

XX Arabidopsis thaliana transcription factor, G308 cDNA.

XX Transcription factor; biochemical characteristic; controlling element;  
DE structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology; plant trait modification; ss.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers

FH CDS 196..1794

FT /\*tag= a

FT /product= "Transcription factor, G308"

XX WO200136597-A1.

PN 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31344.

XX 17-NOV-1999; 99US-0166228.

PR 17-APR-2000; 2000US-0197899.

PR 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (CREE/) CREELMAN R.

PA (YUGG/) YU G.

PA (ADAM/) ADAM L.

PA (RIEC/) RIECHMANN J L.

PA (HEAR/) HEARD J.

PA (SAMA/) SAMAHA R.

PA (PILG/) PILGRIM M.

PA (PINE/) PINEDA O.

PA (JIAN/) JIANG C.

XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;

PI Pilgrim M, Pineda O, Jiang C;

XX WPI: 2001-335999/35.

DR P-PSDB; AAE01907.

CC used to alter the structure and developmental characteristics of plants  
CC such as soybean, wheat, corn, potato, cotton, rice, oilseed rape,  
CC sunflower, alfalfa, sugar cane, turf, banana, blackberry, blueberry,  
CC strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber,  
CC eggplant, grapes, honey dew, lettuce, mango, melon, onion, papaya, peas,  
CC peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato,  
CC watermelon, rosaceous fruits and/or vegetable brassicas. Transcription  
CC factors are key controlling elements of biological pathways and altering  
CC expression levels of 1 or more transcription factors can change entire  
CC biological pathways in an organism. Therefore manipulating transcription  
CC factor levels in plants offers great potential in agricultural  
CC biotechnology for modifying a plant's traits. Transcription factor cDNA  
CC is useful in gene therapy.

XX Sequence: 1951 BP; 482 A; 419 C; 475 G; 575 T; 0 other;

## alignment\_scores:

Quality: 70.00 Length: 17

Ratio: 4.118 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 82.353

## alignment\_block:

US-09-485-529-104 x AAD05791 ..

Align seg 1/1 to: AAD05791 from: 1 to: 1951

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
274 GATGAGCTTCTAGCTGTTCTTGGTTACAAAGTTAGGTCATCGGAATGGC 323

17 a 17

1

324 T 324

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1937.DAT: AAT91937

## seq\_documentation\_block:

ID AAT91937 standard; DNA; 1964 BP.

XX AAT91937;

XX 19-MAR-1998 (first entry)

XX Arabidopsis thaliana gibberellin insensitivty gene gai.

DE Gibberellin insensitivity; gai; plant growth inhibition;  
XX dwarf phenotype; lodging resistance; increased yield;  
KW flowering regulation; bolting inhibition; spinach; lettuce;  
KW antibody; identification; probe; primer; antisense; sense;  
KW expression regulation; co-suppression; rice;  
KW Bakane disease resistance; ss.

XX Arabidopsis thaliana.

XX WO9729123-A2.

PN 14-AUG-1997.

PD 12-FEB-1997; 97WO-GB00390.

XX 12-FEB-1996; 96GB-0002796.

XX (INNE-) INNES CENT INNOVATIONS LTD JOHN.

XX Carol P, Harberd NP, Peng J, Richards DE;

XX WPI: 1997-415295/38.

XX P-PSDB; AAW30792.

XX Nucleic acid encoding gibberellin inhibitor GAI and related

PT antisense sequences - used to create tall, or particularly, dwarf

PT plants, especially crops such as maize, rice and wheat

CC

XX Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the biochemical characteristics of plants e.g. corn,  
PT potato and cotton plants -

PS Claim 4; Page 112-113; 127pp; English.

XX The present sequence is Arabidopsis thaliana transcription factor, G308  
CC cDNA, a homologue of G307. The transcription factor is used for altering  
CC a plant's biochemical characteristics. The transcription factor may be

xx PS Claim 2; Fig 3; 76pp; English.

xx CC The present sequence encodes the Arabidopsis thaliana gibberellin  
xx CC insensitivity (gai) gene product (GAI), the expression of which  
xx CC inhibits plant growth. However the inhibition is antagonised by  
xx CC gibberellin (GA), while gai expression confers a dwarf phenotype  
xx CC that is insensitive to GA. Manipulating gai and GAI expression can  
xx CC produce tall or dwarf plants, particularly the latter for increased  
xx CC resistance to lodging and increased yield. It may also allow  
xx CC regulation of flowering, i.e. plants remain in the vegetative state  
xx CC until treated with GA, useful to inhibit bolting in spinach and  
xx CC lettuce. GAI can be used to raise specific antibodies for  
xx CC identifying homologous proteins or genes in other species. gai  
xx CC fragments can also be used as probes or primers to identify and  
xx CC clone related sequences, or in the preparation of antisense or  
xx CC sense expression regulating (co-suppressing) sequences. Rice plants  
xx CC that express GAI may be resistant to Bakane disease. Manipulation  
xx CC of gai and GAI makes it possible to tailor the degree of dwarfism  
xx CC and GA sensitivity to particular crops or situations.

xx SQ Sequence 1964 BP; 489 A; 426 C; 474 G; 575 T; 0 other;

alignment\_scores:  
Quality: 70.00 Length: 17  
Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x AAT91937 ..

Align seg 1/1 to: AAT91937 from: 1 to: 1964

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
287 GATGAGCTTCTAGCTGTTCTGGTTACAAAGGTAGGTCATCGGAATGCG 336

17 a 17

337 T 337

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF25480

seq\_documentation\_block:

ID AAF25480 standard; DNA; 1779 BP.

XX AC AAF25480;

XX DT 15-MAY-2001 (first entry)

XX DE Nucleotide sequence of a wildtype GRAS protein.

XX KW GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;  
XX KW plant development; dwarf plant; crucifer; ss.

XX OS Brassica napus.

XX FH Key Location/Qualifiers  
XX FT CDS 60..1778  
XX FT /\*tag= a  
XX FT /product= "GRAS"

XX PN WO200109356-A1.

XX XX 08-FEB-2001.

XX XX 02-AUG-2000; 2000WO-FR02216.

XX XX 02-AUG-1999; 99FR-0010023.

XX XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX XX

PI Renard M, Delourme R, Barret P, Brunel D, Froger N, Tanguy X;  
XX DR WPI; 2001-182964/18.  
XX DR P-PSDB; AAB31883.

XX PT New mutant nucleic acid encoding modified GRAS family protein, used to  
XX PT produce dwarf transgenic plants -  
XX XX  
XX PS Example 1; Page 13-15; 28pp; French.

XX CC The present sequence encodes a wild type plant protein of the GRAS  
XX CC family. The specification describes a mutant allele of the BZH gene, the  
XX CC which contains a G1695A mutation resulting in the mutation E546K in the  
XX CC protein. GRAS proteins are transcription factors implicated in  
XX CC regulation of the response to gibberellins and thus in control of  
XX CC morphogenesis and plant development. The mutant GRAS protein is  
XX CC used to produce dwarf plants, specifically crucifers. Dwarf plants may  
XX CC be sown earlier (increasing nitrate accumulation without risking  
XX CC excessive stem growth during winter), and have better resistance to  
XX CC cold and lodging. They are also easier to harvest and allow for better  
XX CC monitoring of the crop.

XX SQ Sequence 1779 BP; 418 A; 430 C; 476 G; 455 T; 0 other;

alignment\_scores:  
Quality: 66.00 Length: 17  
Ratio: 3.882 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

alignment\_block:

US-09-485-529-104 x AAF25480 ..

Align seg 1/1 to: AAF25480 from: 1 to: 1779

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
183 GACGAGCTTCTAGGAGTCTTGGTTACAAAGGTAGGCTTCGGAGATGCG 232

17 a 17

233 T 233

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF25481

seq\_documentation\_block:

ID AAF25481 standard; DNA; 1779 BP.

XX AC AAF25481;

XX DT 15-MAY-2001 (first entry)

XX DE Nucleotide sequence of a mutant GRAS protein.

XX KW GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;  
XX KW plant development; dwarf plant; crucifer; ss.

XX OS Brassica napus.

XX FH Key Location/Qualifiers  
XX FT CDS 60..1778  
XX FT /\*tag= a  
XX FT /product= "GRAS"

XX PN WO200109356-A1.

XX XX 08-FEB-2001.

XX XX 02-AUG-2000; 2000WO-FR02216.

XX XX 02-AUG-1999; 99FR-0010023.

XX XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX PA

XX Renard M, Delourme R, Barret P, Brunel D, Froger N, Tanguy X;  
 XX WPI; 2001-182964/18.  
 DR P-PSDB; AAB31884.  
 XX  
 XX New mutant nucleic acid encoding modified GRAS family protein, used to  
 PT produce dwarf transgenic plants -  
 PS Example 1; Page 18-20; 28pp; French.  
 XX  
 CC The present sequence encodes a mutant plant protein of the GRAS  
 CC family. The mutant allele of the B2H gene contains a G1695A mutation  
 CC resulting in the mutation E546K in the protein. GRAS proteins are  
 CC transcription factors implicated in regulation of the response to  
 CC gibberellins and thus in control of morphogenesis and plant development.  
 CC The mutant GRAS protein is used to produce dwarf plants, specifically  
 CC crucifers. Dwarf plants may be sown earlier (increasing nitrate  
 CC accumulation without risking excessive stem growth during winter), and  
 CC have better resistance to cold and lodging. They are also easier to  
 CC harvest and allow for better monitoring of the crop.  
 XX  
 SQ Sequence 1779 BP; 419 A; 430 C; 475 G; 455 T; 0 other;

alignment\_scores:  
 Quality: 66.00 Length: 17  
 Ratio: 3.882 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 76.471  
 alignment\_block:  
 US-09-485-529-104 x AAF25481 ..  
 Align seg 1/1 to: AAF25481 from: 1 to: 1779  
 1 AspGluLeuLeuAlaLaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||.....:|||||.....:|||||.....:|||||  
 183 GACGAGCTCTAGGAGTCTTGGGTACAAAGGTAGGTCTTCGGAGATGCG 232  
 17 a 17  
 |  
 233 T 233  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AA36266  
 seq\_documentation\_block:  
 ID AAX36266 standard; DNA; 200 BP.  
 XX  
 AC AAX36266;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE DNA sequence obtained after sequencing wheat Rht clone 14a1.  
 XX  
 KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 PN WO9909174-A1.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 07-AUG-1998; 98WO-GB02383.  
 XX  
 PR 13-AUG-1997; 97GB-0017192.  
 XX  
 PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX  
 PI Harberd NP, Peng J, Richards DE;  
 XX  
 XX WPI; 1999-181040/15.  
 DR

XX New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype  
 XX  
 PS Disclosure; Fig 2b(12); 88pp; English.  
 XX  
 CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in Triticum Aestivum, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence was  
 CC obtained after partially sequencing wheat Rht clone 14a1.  
 XX  
 SQ Sequence 200 BP; 24 A; 58 C; 84 G; 27 T; 7 other;

alignment\_scores:  
 Quality: 65.00 Length: 15  
 Ratio: 4.643 Gaps: 0  
 Percent Similarity: 93.333 Percent Identity: 93.333  
 alignment\_block:  
 US-09-485-529-104 x AAX36266 ..  
 Align seg 1/1 to: AAX36266 from: 1 to: 200  
 1 AspGluLeuLeuAlaLaLeuGlyTyrLysValArgAlaSerAsp 15  
 |||||.....:|||||.....:|||||.....:|||||  
 79 GACTATCTGCTGCGCGCTCGGTACAAAGGTGCGCGCTCCGAC 123  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AA36283  
 seq\_documentation\_block:  
 ID AAX36283 standard; DNA; 416 BP.  
 XX  
 AC AAX36283;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE Partial sequence of the wheat rht-10 allele.  
 XX  
 KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 PN WO9909174-A1.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 07-AUG-1998; 98WO-GB02383.  
 XX  
 PR 13-AUG-1997; 97GB-0017192.  
 XX  
 PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX  
 PI Harberd NP, Peng J, Richards DE;  
 XX  
 DR WPI; 1999-181040/15.  
 DR P-PSDB; AAY02544.  
 XX  
 PT New Triticum Aestivum polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is

PT antagonised by gibberellin, used to confer a dwarf phenotype  
 PS Disclosure; Fig 12a; 88pp; English.

CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in *Triticum aestivum*, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence represents  
 CC the partial sequence of the wheat rht-10 allele.

SQ Sequence 416 BP; 59 A; 168 C; 129 G; 60 T; 0 other;

alignment\_scores:  
 Quality: 62.00 Length: 13  
 Ratio: 4.769 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-485-529-104 x AAX36283 ..

Align seg 1/1 to: AAX36283 from: 1 to: 416

5 AlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 |||||  
 1 GCGGCGCTGGGTACAAAGGTGCGCGCTCGACATGGCG 39

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX36261

seq\_documentation\_block:

ID AAX36261 standard; DNA; 309 BP.

AC AAX36261;

DT 16-JUL-1999 (first entry)

DE DNA sequence obtained after sequencing wheat Rht clone 14a1.

KW Rht gene; homologue; *Triticum aestivum*; wheat; growth inhibition;  
 KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
 KW paclobutrazol; ss.

OS *Triticum aestivum*.

XX WO9909174-A1.

PD 25-FEB-1999.

PF 07-AUG-1998; 98WO-GB02383.

PR 13-AUG-1997; 97GB-0017192.

PA (PLAN-) PLANT BIOSCIENCE LTD.

PI Harberd NP, Peng J, Richards DE;

DR WPI; 1999-181040/15.

PT New *Triticum aestivum* polynucleotides - encode a polypeptide which  
 PT provides inhibition of the growth of plants, which inhibition is  
 PT antagonised by gibberellin, used to confer a dwarf phenotype

PS Disclosure; Fig 2b(7); 88pp; English.

CC The specification describes polypeptides encoded by the Rht gene (and  
 CC its homologues) that, when expressed in *Triticum aestivum*, inhibit  
 CC growth of the plant. This growth inhibition is antagonised by  
 CC gibberellin. The products can be used to provide Rht expression in  
 CC plants, conferring a dwarf phenotype on a plant which is correctable  
 CC by treatment with gibberellin. In addition, the products can be  
 CC used to produce Rht mutant plants which are dwarfed compared with  
 CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
 CC may be made by knocking out Rht or the relevant homologous gene in  
 CC the plant of interest. Plants may be made which are resistant to  
 CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
 CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
 CC dwarf but let crop plants grow tall. The present sequence was  
 CC obtained after partially sequencing wheat Rht clone 14a1.

SQ Sequence 309 BP; 47 A; 102 C; 102 G; 45 T; 13 other;

alignment\_scores:  
 Quality: 46.00 Length: 9  
 Ratio: 5.111 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-485-529-104 x AAX36261 ..

Align seg 1/1 to: AAX36261 from: 1 to: 309

8 GlyTyrLysValArgAlaSerAspMet 16  
 |||||  
 18 GGGTACAAGGTGCGCGCTCCGACATG 44

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT: AAT86824

seq\_documentation\_block:

ID AAT86824 standard; cDNA; 817 BP.

AC AAT86824;

DT 09-JAN-1998 (first entry)

DE Hevea brasiliensis (S)-hydroxynitrilase cDNA.

KW (S)-hydroxynitrilase; production; (S)-cyanohydrin; aldehyde;  
 KW hydrogen cyanide; recombinant; ss.

OS Hevea brasiliensis.

FH Key Location/Qualifiers  
 FT CDS 44..817  
 FT /\*tag= a  
 FT /product= (S)-hydroxynitrilase

XX DE19529116-A1.

XX 06-MAR-1997.

XX 08-AUG-1995; 95DE-1029116.

XX 08-AUG-1995; 95DE-1029116.

XX (OSTS ) CHEMIE LINZ DEUT GMBH IL.

XX WPI; 1997-155423/15.

XX P-PSDB; AAW29164.

PT DNA encoding Hevea brasiliensis (S)-hydroxynitrilase - useful for  
 PT production of cyanohydrin(s) from aldehyde and hydrogen cyanide

PS Claim 2; Page 8; 8pp; German.

CC The present sequence encodes Hevea brasiliensis

CC (S)-hydroxynitrilase (S-HN), useful in the production of

CC (S)-cyanohydrins from aldehyde and hydrogen cyanide. Recombinantly  
 CC produced S-HN has higher specific activity than native S-HN,  
 CC probably due to post translational modification differences  
 CC between plants and microorganisms.  
 CC S-HN was isolated from H. brasiliensis leaves by homogenisation,  
 CC followed by sequential chromatography on QAE-Sepharose FF,  
 CC Phenyl-Sepharose and BioGel P180. A cDNA bank from the same source  
 CC was prepared and screened with polyclonal rabbit antiserum raised  
 CC against the isolated enzyme. The insert, about 1.1 kb, in the  
 CC single positive clone was cloned into pHLN100 and sequenced.  
 XX  
 SQ Sequence 817 BP; 250 A; 159 C; 193 G; 215 T; 0 other;

alignment\_scores:  
 Quality: 45.00 Length: 15  
 Ratio: 3.750 Gaps: 0  
 Percent Similarity: 80.000 Percent Identity: 66.667

alignment\_block:

US-09-485-529-104 x AAT86824 ..

Align seg 1/1 to: AAT86824 from: 1 to: 817

3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 ||||| |||||:::||||| || |||||:::|||||  
 116 CTCCTGTGAGGCACCTGGCCACAGCTTACTGCACCTGGACCTTGCA 160

seq\_name: /SDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV36351

seq\_documentation\_block:

ID AAT36351 standard; cDNA; 1091 BP.

XX AAT36351;

XX 29-OCT-1997 (first entry)

DE Hevea brasiliensis S-hydroxynitrilase cDNA.

KW S-hydroxynitrilase; chiral synthesis; aliphatic; aromatic;

KW S-cyanohydrin; recombinant; ss.

XX Hevea brasiliensis.

XX Key Location/Qualifiers

FT CDS 57..830

FT /\*tag= a

FT /product= S-hydroxynitrilase

XX WO9703204-A2.

XX 30-JAN-1997.

XX 10-JUL-1996; 96WO-EP03010.

XX 12-JUL-1995; 95AT-0001182.

XX (STAM ) DSM CHEMIE LINZ GMBH.

XX Griengl H, Hasslacher M, Hayn EM, Kohlwein S, Schall M;  
 PI Schwab H;

XX WPI; 1997-119058/11.

XX P-PSDB; AAW10593.

XX Hevea brasiliensis S-hydroxynitrilase - and recombinant equivalent  
 PT of high specific activity, for chiral synthesis of cyano:hydrin(s)

XX Claim 2; Pages 2-3; 14pp; German.

XX The present sequence encodes the Hevea brasiliensis

CC S-hydroxynitrilase (S-HN), which can be used for the chiral

CC synthesis of aliphatic or aromatic S-cyanohydrins from HCN, or its

CC donor, and ketone. Recombinant S-HN has higher specific activity  
 CC than the native enzyme.  
 CC H. brasiliensis leaves were homogenised in cold pH 6.5 K phosphate  
 CC buffer, and the extract subjected to sequential chromatography on  
 CC QAE-Sepharose, Phenyl-Sepharose and BioGel 150 to recover a protein  
 CC with a specific activity of 19 IU/mg.  
 CC A cDNA library was prepared from young H. brasiliensis leaves in  
 CC Zap phage, and screened with polyclonal rabbit antiserum raised  
 CC against the purified enzyme. The insert in one positive clone was  
 CC isolated, and ligated into pBluescript to form pHLN-100. This  
 CC plasmid, in E. coli SOLR, expressed a fusion protein with lacZ that  
 CC was immunoreactive. Sequencing showed that the insert in this  
 CC plasmid was not complete, and the missing 5'-end generated by 2  
 CC stage PCR, resulting in plasmid pHLN-101, containing the entire  
 CC coding sequence.  
 XX  
 SQ Sequence 1091 BP; 326 A; 196 C; 239 G; 330 T; 0 other;

alignment\_scores:  
 Quality: 45.00 Length: 15  
 Ratio: 3.750 Gaps: 0  
 Percent Similarity: 80.000 Percent Identity: 66.667

alignment\_block:

US-09-485-529-104 x AAT36351 ..

Align seg 1/1 to: AAT36351 from: 1 to: 1091

3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 ||||| |||||:::||||| || |||||:::|||||  
 129 CTCCTGTGAGGCACCTGGCCACAGCTTACTGCACCTGGACCTTGCA 173

seq\_name: /SDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV36418

seq\_documentation\_block:

ID AAV36418 standard; cDNA; 1458 BP.

XX AAV36418;

XX 14-SEP-1998 (first entry)

DE Keratan sulphate 6-sulphotransferase.

KW K5Gal6ST; keratan sulphate 6-sulphotransferase; hybridization;

KW chick chondroitin 6-sulphotransferase; C6ST; phage lambda;

KW expression vector; COS-7 cells; Bluescript plasmid; galactose;

XX keratan sulphate; chondroitin; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1236

FT /\*tag= a

FT /product= "Keratan sulphate 6-sulphotransferase"

XX EP845533-A2.

XX 03-JUN-1998.

XX 27-NOV-1997; 97EP-0309564.

XX 29-NOV-1996; 96JP-0320535.

XX (SEKG ) SEIKAGAKU KOGYO CO LTD.

XX Fukuta M, Habuchi O;

XX WPI; 1998-288750/26.

XX P-PSDB; AAW61100.

XX Keratan sulphate 6-sulpho-transferase polypeptide - transfers  
 PT sulphate from sulphate donor to galactose 6-hydroxy group etc.









```
alignment_scores:
  Quality: 43.00      Length: 9
  Ratio: 4.778      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-09-485-529-104 x AAX36248 ..
Align seg 1/1 to: AAX36248 from: 1 to: 29

      8 GlyTyrLysValArgAlaSerAspMet 16
      |||||||:|||||:|||||
      3 GGGTACAGAGTGGTTCGCGATG 29

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2000.DAT:AAC47519

seq_documentation_block:
ID AAC47519 standard; DNA; 864 BP.
XX
AC AAC47519;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 54118.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
PF 06-SEP-2000.
XX
PP
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
```



PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

CC This sequence encodes a H. pylori inner membrane protein.  
CC Helicobacter pylori has been strongly linked to chronic gastritis and  
CC duodenal ulcer disease. The nucleic acid sequences of the invention  
CC are used to evaluate compounds, especially activators or inhibitors of  
CC bacterial life cycle, for the ability to bind an H. pylori nucleic acid  
CC sequence. The nucleic acid sequences, and corresponding proteins, are  
CC also useful for generating vaccines for immunising subjects against H.





```
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157533.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 42.00      Length: 15
  Ratio: 3.500      Gaps: 0
Percent Similarity: 80.000 Percent Identity: 60.000

alignment_block:
US-09-485-529-104 x AAC47916 ..
Align seg 1/1 to: AAC47916 from: 1 to: 1325
3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetala 17
98 CTTCTCGAGGCTTTGGCCCTGTAAACGCCCTTAGACCTAGCT 142

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAx30475
seq_documentation_block:
ID AAX30475 standard; DNA; 1464 BP.
XX
AC AAX30475;
XX
DT 08-JUN-1999 (first entry)
XX
DE H. pylori secreted protein ORF hp6p10606_23493756.cl_21.
XX
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
secreted protein; cellular protein; ds.
XX
OS Helicobacter pylori.
XX
PN WO9818323-Al.
XX
PD 07-MAY-1998.
XX
PF 28-OCT-1997; 97WO-US19575.
XX
PR 14-JUL-1997; 97US-0891928.
PR 28-OCT-1996; 96US-0739150.
PR 06-DEC-1996; 96US-0759739.
XX
PA (ASTR ) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1998-271811/24.
DR P-PSDB; AAY11008.
XX
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections
XX
PS Claims 3, 4; Page 137; 279pp; English.
XX
CC Recombinant or substantially pure preparations of H. pylori polypeptides
are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
```

CC The protein may be used in a vaccine to prevent or treat *H. pylori*





Quality: 41.00 Length: 14  
Ratio: 3.727 Gaps: 0  
Percent Similarity: 78.571 Percent Identity: 64.286

alignment\_block:

US-09-485-529-104 x AAZ10093

Align seg 1/1 to: AAZ10093 from: 1 to: 235

2 GluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15  
||||||| :||| ||||| ||||| ||||| |||||  
84 GAATTATGAACGTCAAGGTTATCTCTGTTAAAGCATCCGAT 125

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ10093

seq\_documentation\_block:

ID AAZ36271 standard; DNA; 285 BP.

XX AC AAZ36271;

DT 16-JUL-1999 (first entry)

DE DNA sequence obtained after sequencing wheat Rht clone 5a1.

KW Rht gene; homologue; Triticum aestivum; wheat; growth inhibition;  
KW antagonist; gibberellin; dwarf phenotype; gibberellin biosynthesis;  
KW paclobutrazol; ss.

OS Triticum aestivum.

XX WO9909174-A1.

PN 25-FEB-1999.

PD 07-AUG-1998; 98WO-GB02383.

PP 13-AUG-1997; 97GB-0017192.

PR (PLAN-) PLANT BIOSCIENCE LTD.

PI Harberd NP, Peng J, Richards DE;

XX WPI; 1999-181040/15.

DR New Triticum Aestivum polynucleotides - encode a polypeptide which  
PT provides inhibition of the growth of plants, which inhibition is  
PT antagonised by gibberellin, used to confer a dwarf phenotype

XX Disclosure: Fig 2c(4); 88pp; English.

XX The specification describes polypeptides encoded by the Rht gene (and  
CC its homologues) that, when expressed in Triticum aestivum, inhibit  
CC growth of the plant. This growth inhibition is antagonised by  
CC gibberellin. The products can be used to provide Rht expression in  
CC plants, conferring a dwarf phenotype on a plant which is correctable  
CC by treatment with gibberellin. In addition, the products can be  
CC used to produce Rht mutant plants which are dwarfed compared with  
CC wild-type, the dwarfing being gibberellin-insensitive. Taller plants  
CC may be made by knocking out Rht or the relevant homologous gene in  
CC the plant of interest. Plants may be made which are resistant to  
CC compounds which inhibit gibberellin biosynthesis, such as paclobutrazol,  
CC e.g. to allow use of a gibberellin biosynthesis inhibitor to keep weeds  
CC dwarf but let crop plants grow tall. The present sequence was  
CC obtained after partially sequencing wheat Rht clone 5a1.

XX Sequence 285 BP; 39 A; 89 C; 100 G; 46 T; 11 other;

alignment\_scores:

Quality: 41.00 Length: 9  
Ratio: 5.125 Gaps: 0  
Percent Similarity: 88.889 Percent Identity: 88.889

alignment\_block:

US-09-485-529-104 x AAX36271/rev

Align seg 1/1 to reverse of: AAX36271 from: 1 to: 285

8 GlyTyrLysValArgAlaSerAspMet 16  
||||||| ||||| ||||| ||||| |||||  
283 GGGTACAGGNGCGGCATCGACATG 257

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH99418

seq\_documentation\_block:

ID AAH99418 standard; cDNA; 556 BP.

XX AC AAH99418;

XX 16-OCT-2001 (first entry)

DT Human protein encoding cDNA sequence SEQ ID NO:253.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiagregant; haemostatic; vulnery; antidiabetic; cytostatic;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.

XX Homo sapiens.

XX WO200153455-A2.

XX 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US35017.

XX 23-DEC-1999; 99US-0471275.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX P-PSDB; AAM25477.

XX Isolated human polynucleotides encoding polypeptides, useful for the  
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX Claim 1; Page 410; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
XX AAM25963. The proteins can have activities based on the tissues and  
XX cells they are expressed in, such as: antiinflammatory; antirheumatic;  
XX antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
XX central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
XX cardiovascular; antianaemic; antiagregant; haemostatic; vulnery;  
XX antidiabetic; cytostatic; dermatological; antiallergic; antiasthmatic;  
XX antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
XX encoding them can be used in gene therapy, antisense therapy and vaccine  
XX production, the proteins and polynucleotides are useful for screening for  
XX agonists or antagonists of a protein and for the treatment and diagnosis  
XX of disorders associated with the activity of a protein e.g. inflammation,

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, parkinson's disease, neurodegenerative and  
 CC neurological disorders.  
 XX  
 SQ Sequence 556 BP; 153 A; 125 C; 128 G; 150 T; 0 other;  
 alignment\_scores:  
 Quality: 41.00 Length: 14  
 Ratio: 3.154 Gaps: 0  
 Percent Similarity: 92.857 Percent Identity: 50.000  
 alignment\_block:  
 US-09-485-529-104 x AAH99418/rev ..  
 Align seg 1/1 to reverse of: AAH99418 from: 1 to: 556  
 3 LeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMet 16  
 ::::|||||:|||||:|||||:|||||:|||||:|  
 294 ATGATGGCAGCATTTGGGATACAGTGTGAAAAGCAGCACATTG 253  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH81463  
 seq\_documentation\_block:  
 ID AAH81463 standard; DNA; 1476 BP.  
 XX  
 AC AAH81463;  
 XX  
 DT 21-SEP-2001 (first entry)  
 XX  
 DE Escherichia coli protein encoding nucleotide sequence SEQ ID NO:262.  
 XX  
 KW Escherichia coli; identification; proliferation; microorganism;  
 KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;  
 KW bacterial growth inhibition; ds.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO200148209-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 19-DEC-2000; 2000WO-US34419.  
 XX  
 PR 23-DEC-1999; 99US-0173005.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Forsyth RA, Ohlsen KL, Zyskind JW;  
 XX  
 DR WPI: 2001-457376/49.  
 DR P-PSDB; AAG98407.  
 XX  
 PT Novel nucleic acids encoding proteins required for Escherichia coli  
 PT proliferation, useful for screening for antimicrobial agents -  
 XX  
 PS Claim 9; Page 393-395; 596pp; English.  
 XX  
 CC The present invention describes a purified or isolated nucleic acid  
 CC sequence (I) consisting essentially of one of the 93 nucleotide sequences  
 CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a  
 CC microorganism is capable of inhibiting proliferation of a microorganism.  
 CC (I) have antibacterial and antibiotic activities, and can be used in  
 CC gene therapy. Expression of (I) in a microorganism inhibits proliferation  
 CC of the microorganism, and the manufactured antibiotic is useful for  
 CC reducing the activity or level of a gene product required for  
 CC proliferation of a microorganism in a subject, specifically humans. The  
 CC nucleic acids that inhibit bacterial growth or proliferation can be used

CC as antisense therapeutics for killing bacteria. In addition to  
 CC therapeutic applications, the nucleic acid sequences complementary to  
 CC sequences required for proliferation can be used as diagnostic tools.  
 CC For example, nucleic acid probes complementary to proliferation-required  
 CC sequences that are specific for particular species of microorganisms can  
 CC be used as probes to identify particular microorganism species in  
 CC clinical specimens. AAH81295 to AAH81487 encode the Escherichia coli  
 CC proteins given in AAG98239 to AAG98431, and AAH81488 to AAH81491  
 CC represent oligonucleotides, which are used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 1476 BP; 330 A; 372 C; 436 G; 338 T; 0 other;  
 alignment\_scores:  
 Quality: 41.00 Length: 16  
 Ratio: 3.154 Gaps: 0  
 Percent Similarity: 81.250 Percent Identity: 50.000  
 alignment\_block:  
 US-09-485-529-104 x AAH81463 ..  
 Align seg 1/1 to: AAH81463 from: 1 to: 1476  
 2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 :|||||:|||||:|||||:|||||:|||||:|  
 109 GAAGTTCGGCCAAATGAAGGTATCAGATCAGTGGTCCGATTTAGCG 156  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ53179  
 seq\_documentation\_block:  
 ID AAZ53179 standard; DNA; 1521 BP.  
 XX  
 AC AAZ53179;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE Neisseria gonorrhoeae ORF 092 partial DNA sequence SEQ ID NO:309.  
 XX  
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy; ds.  
 XX  
 OS Neisseria gonorrhoeae.  
 XX  
 PN WO9957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 25-FEB-1999; 98US-0121528.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;  
 PI Petersen J, Piazza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI: 2000-062150/05.  
 DR P-PSDB; AAY74417.  
 XX  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 XX

```

PS Claim 7; Page 288; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615, and AA254615 to AA254615
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254616 to AA254616 to AA254616
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria meningitidis (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria meningitidis, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 1521 BP; 359 A; 430 C; 394 G; 338 T; 0 other;

alignment_scores:
  Quality: 41.00 Length: 16
  Ratio: 3.154 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 56.250

alignment_block:
US-09-485-529-104 x AA253179 ..
Align seg 1/1 to: AA253179 from: 1 to: 1521

2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerMetala 17
||||:|||| |:::::||||| :||||:|||||
181 GAAGTCTGCACAAATTGGGCTTTAAAGTTTCGGTTCGGATCAGCGC 228

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ53180
seq_documentation_block:
ID AAZ53180 standard; DNA; 1521 BP.
XX
AC AAZ53181;
XX
DT 21-WAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 092 partial DNA sequence SEQ ID NO:311.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.

DR P-PSDB; AA254418.
XX
PT Novel Neisseria meningitidis predicted to be useful antigens for
XX vaccines and diagnostics
XX
PS Claim 7; Page 289; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615, and AA254615 to AA254615
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254616 to AA254616 to AA254616
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria meningitidis (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria meningitidis, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 1521 BP; 352 A; 444 C; 396 G; 329 T; 0 other;

alignment_scores:
  Quality: 41.00 Length: 16
  Ratio: 3.154 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 56.250

alignment_block:
US-09-485-529-104 x AA253180 ..
Align seg 1/1 to: AA253180 from: 1 to: 1521

2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerMetala 17
||||:|||| |:::::||||| :||||:|||||
181 GAAGTCTGCACAAATTGGGCTTTAAAGTTTCGGTTCGGATCAGCGC 228

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ53181
seq_documentation_block:
ID AAZ53181 standard; DNA; 1521 BP.
XX
AC AAZ53181;
XX
DT 21-WAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 092 partial DNA sequence SEQ ID NO:313.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX

```

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI: 2000-062150/05.  
 DR P-PSDB; AAY74419.  
 XX  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 PT  
 PS Claim 7; Page 290-291; 1453pp; English.  
 XX  
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAY75941  
 CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 1521 BP; 353 A; 436 C; 397 G; 335 T; 0 other;

alignment\_scores:  
 Quality: 41.00 Length: 16  
 Ratio: 3.154 Gaps: 0  
 Percent Similarity: 81.250 Percent Identity: 56.250  
 alignment\_block:  
 US-09-485-529-104 x AAZ53181 ..  
 Align seg 1/1 to: AAZ53181 from: 1 to: 1521  
 2 GluteuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 |||::||| |||::||| |||::||| |||::||| ::|||::|||  
 181 GAAGTCGTGCACATTTGGGTTTAAAGTTTCCGGTTCGGATCAGGCG 228

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ71977  
 seq\_documentation\_block:  
 ID AAQ71977 standard; DNA; 1608 BP.

XX AAQ71977;  
 XX  
 XX 03-MAY-1995 (first entry)  
 XX  
 DE Murine IL-2R gamma gene.  
 XX  
 KW Murine IL-2R gamma; X-linked severe combined immunodeficiency;  
 KW XSCID; interleukin; ss.  
 XX  
 OS Mus musculus.  
 XX  
 XX Key Location/Qualifiers  
 FH 25..1134  
 FT CDS /\*tag= a  
 FT /\*transl\_except= pos:1015..1017, aa:His  
 FT sig\_peptide 25..82  
 FT /\*tag= b  
 FT /\*note= "Putative"  
 FT  
 XX WO9420641-A.  
 XX  
 PN 15-SEP-1994.  
 XX  
 PD 10-MAR-1994; 94WO-US02891.  
 XX

PR 12-MAR-1993; 93US-0031143.  
 PR 14-SEP-1993; 93US-0121435.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Leonard WJ, McBride WO, Noguchi M;  
 XX  
 DR WPI; 1994-303046/37.  
 DR P-PSDB; AAR59094.  
 XX  
 PT Diagnosis of X-linked severe combined immunodeficiency (XSCID) -  
 PT comprises detecting mutated IL-2R gamma gene, also vectors and  
 PT transgenic animals containing the mutated gene  
 XX  
 PS Example 1; Fig 7; 98pp; English.  
 XX  
 CC AAQ71977 is the DNA sequence of murine IL-2R gamma AAR59094,  
 CC this was used in the development of a claimed method for the  
 CC diagnosis of X-linked severe combined immunodeficiency (XSCID),  
 CC in female carriers and male sufferers.  
 XX  
 SQ Sequence 1608 BP; 413 A; 434 C; 349 G; 412 T; 0 other;

alignment\_scores:  
 Quality: 41.00 Length: 13  
 Ratio: 3.727 Gaps: 0  
 Percent Similarity: 84.615 Percent Identity: 69.231

alignment\_block:  
 US-09-485-529-104 x AAQ71977/rev ..  
 Align seg 1/1 to reverse of: AAQ71977 from: 1 to: 1608

2 GluteuLeuAlaAlaLeuGlyTyrLysValArgAlaSer 14  
 ||||| |||::|||::|||::|||::|||::|||  
 1409 GAGCTCCAGCACCCCTGGCTACAGGTAGACTCTCT 1371  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS06754  
 seq\_documentation\_block:  
 ID AAS06754 standard; cDNA; 2730 BP.

XX AAS06754;  
 XX  
 XX 12-SEP-2001 (first entry)  
 XX  
 DE Polynucleotide sequence encoding human protein kinase #54.  
 XX  
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200138503-A2.  
 XX  
 PD 31-MAY-2001.  
 XX  
 PF 22-NOV-2000; 2000WO-US32085.  
 XX  
 PR 24-NOV-1999; 99US-0167482.  
 XX  
 PA (SUGE-) SUGEN INC.  
 XX  
 PI Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PI Flanagan P, Clary D;  
 XX  
 DR WPI; 2001-343950/36.  
 DR P-PSDB; AAU03554.  
 XX  
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing









gb_estl1:B5629895	+	40.00	125.69	775.41	401	1 BEG29895 uu36c09.x1 Soares_mamm	gb_gss:CNS074PB8	-	40.00	118.71	1.9e+03	879	1 AL429090 clone BA0AB029F07 o
gb_est2:B3660695	+	40.00	125.67	777.62	402	1 BG52985 TgEsty5a15e10.y1 TgVPC	gb_est2:BF232620	-	40.00	118.54	1.9e+03	896	1 BF232620 602023203F1 NCI CGA
gb_est2:B3684221	+	40.00	125.47	797.50	411	1 BG984221 PM0-CN0150-200301-008	gb_est2:BG527856	-	40.00	118.49	2.0e+03	901	1 BG527856 602556813F1 NIH MGC
gb_est1:AA142418	+	40.00	125.43	801.93	413	1 AA142418 ms08h04.r1 Stratagene	gb_est2:BG969195	-	40.00	118.39	2.0e+03	911	1 BE969195 601649401R1 NIH MGC
gb_est2:W96848	+	40.00	125.43	801.93	413	1 W96848 mif5c05.r1 Soares mouse	gb_gss:CNS06E9D	-	40.00	118.05	2.1e+03	947	1 AL394823 T7 end of clone ARO
gb_est1:AV162063	+	40.00	125.40	804.14	414	1 AV162063 AV162063 MS musculus	gb_gss:CNS076N0	-	40.00	117.78	2.1e+03	976	1 AL394823 T7 end of clone ARO
gb_est1:A1772824	+	40.00	125.28	817.45	420	1 A1772824 EST253924 Tomato resis	gb_gss:CNS05IN58	-	40.00	116.78	2.4e+03	1032	1 AK039127 Tetraodon nigrovir
gb_est2:BG660362	+	40.00	125.21	824.11	423	1 BG660362 TgEsty2y07a02.y1 TgVPC	gb_hlc:AK0031958	-	40.00	114.83	3.1e+03	1359	1 AK039127 Mus musculus 18 da
gb_est2:BF557639	+	40.00	125.17	825.56	425	1 BF557639 UI-R-CO-bx-a-07-0-UI-R	gb_hlc:AK008870	-	40.00	112.31	4.3e+03	1806	1 AK008870 Mus musculus adult
gb_est2:T27048	+	40.00	124.92	858.29	437	1 T27048 NIBT219C09R Infant brain	gb_est2:BF785554	-	39.50	116.63	2.5e+03	923	1 BF785554 602112058F1 NCI CGA
gb_est2:BG393690	+	40.00	124.84	864.22	441	1 BG393690 OVL1_L6.D06.B1.A002 Ova	gb_est2:BA080759	-	39.00	128.41	546.77	204	1 BA080759 0190F Pyrococcus fu
gb_est2:BG993104	-	40.00	124.78	870.93	444	1 BG993104 RC4-H51109-0620201-014	gb_est1:AA1436954	-	39.00	128.24	559.02	208	1 AA1436954 fb36c01.y1 Zebrafis
gb_gss:AO682893	-	40.00	124.74	875.41	446	1 AO682893 HS-2229.B1.C03 SP6E RF	gb_est1:AW121277	-	39.00	126.75	676.91	246	1 AW121277 UI-M-BH2.3-rog-b-09
gb_est1:BE162592	+	40.00	124.70	879.89	448	1 BE162592 PM1-HT0452-150300-004	gb_est1:AW404912	-	39.00	126.64	676.91	246	1 AW404912 UI-HF-BL0-acn-h-02
gb_est2:T86731	+	40.00	124.60	891.09	453	1 T86731 yd77f05.x1 Soares fetal	gb_est1:AV257911	-	39.00	126.53	695.77	252	1 AV257911 AV257911 RIKEN full
gb_est2:BF151780	+	40.00	124.60	891.09	453	1 BF151780 uz20c11.y1 NCI CGAP_Ma	gb_est1:AV257911	-	39.00	126.53	695.77	252	1 AV257911 AV257911 RIKEN full
gb_gss:AZA2421997	+	40.00	124.43	911.31	462	1 AZA2421997 IM0200E08 Mouse 10kb	gb_gss:BH035748	-	39.00	126.50	698.92	253	1 BH035748 RPI2-24-387F4.IVB R
gb_est2:BG904251	+	40.00	124.28	929.33	470	1 BG904251 TaLR1103F05F TaLR1 Tr1	gb_est2:BF050022	-	39.00	126.09	736.85	265	1 BF050022 CM2-GN0294-020101-6
gb_est1:AW824521	+	40.00	124.22	936.10	473	1 AW824521 us10612.x1 Soares NMGE	gb_est1:AA917391	-	39.00	125.99	745.37	268	1 AA917391 O180806.s1 NCI CGAP
gb_gss:AZ825888	-	40.00	124.05	956.44	482	1 AZ825888 2M0101N13F Mouse 10kb	gb_est2:D24394	-	39.00	125.69	775.03	277	1 D24394 RICR1832A Rice root O
gb_gss:AO502963	-	40.00	123.92	972.30	489	1 AO502963 V47C6 mfn-3xHA/lacZ In	gb_est1:BB286985	-	39.00	125.31	813.43	289	1 BB286985 BB286985 RIKEN full
gb_est1:AA646019	-	40.00	123.89	976.83	491	1 AA646019 vS29412.r1 Stratagene	gb_est1:BB286985	-	39.00	125.31	813.43	289	1 BB286985 BB286985 RIKEN full
gb_est1:AA880174	-	40.00	123.85	981.37	493	1 AA880174 v97g06.r1 Soares_mamm	gb_est1:BB516804	-	39.00	125.28	816.64	290	1 BB516804 BB516804 RIKEN full
gb_est1:AA637496	-	40.00	123.82	985.91	495	1 AA637496 vr32c05.r1 Barstead mc	gb_gss:AO605631	-	39.00	125.07	839.16	297	1 AO605631 HS-5386.A1.D10.MR R
gb_est1:AA270703	-	40.00	123.78	990.46	497	1 AA270703 va67b05.r1 Soares mous	gb_gss:BE706303	-	39.00	124.92	855.29	302	1 BE706303 RCI-HT0256-150600-1
gb_est1:AW908843	-	40.00	123.55	1.0e+03	510	1 AW908843 uf38c01.x1 Soares_mamm	gb_gss:AZ285604	-	39.00	124.69	881.18	310	1 AZ285604 RPI2-23-155G3.TV RP
gb_est1:A1649652	-	40.00	123.46	1.0e+03	515	1 A1649652 48607C12.y1 486 - lea	gb_est1:AW851095	-	39.00	124.66	884.42	311	1 AW851095 IL3-CT0220-150200-1
gb_gss:TA137804P	+	40.00	123.33	1.0e+03	523	1 AL65954 T. brucei sheared genc	gb_est2:BE933120	-	39.00	124.63	887.67	312	1 BE933120 RCI-HT0256-240800-1
gb_est1:AA685654	-	40.00	123.28	1.1e+03	526	1 AA685654 vufb12.r1 Barstead mc	gb_gss:CNS00W08	-	39.00	124.44	910.41	319	1 CNS00W08 Arabidopsis thalian
gb_est1:AW753318	-	40.00	123.19	1.1e+03	531	1 AW753318 RC5-CT0254-031099-011	gb_est2:BE599140	-	39.00	124.38	916.93	321	1 BE599140 EST504004 csts Sola
gb_est1:BB433359	-	40.00	123.12	1.1e+03	535	1 BB433359 EST399888 Tomato break	gb_est1:AA413598	-	39.00	124.33	923.44	323	1 AA413598 vc56d01.s1 Knowles
gb_est2:BF470480	-	40.00	123.11	1.1e+03	536	1 BF470480 UI-M-BH3-avl-h-12-0-UI	gb_gss:AO721579	-	39.00	124.13	946.30	330	1 AO721579 HS-5557.A1.E11.SP6E
gb_est1:A1875584	+	40.00	123.09	1.1e+03	537	1 A1875584 uk50b05.x1 Sugano mous	gb_est1:AW175826	-	39.00	124.11	949.57	331	1 AW175826 CM0-TF0057-260899-0
gb_est1:AW744749	+	40.00	123.03	1.1e+03	541	1 AW744749 ur33a08.x1 Soares_mous	gb_est1:BE706286	-	39.00	124.03	959.40	334	1 BE706286 RCI-HT0256-150600-1
gb_est1:BE626837	+	40.00	123.01	1.1e+03	542	1 BE626837 ut85a05.x1 Soares_mamm	gb_est1:AO32548	-	39.00	123.90	975.79	339	1 AO32548 AV032548 Mus muscul
gb_est1:BE750188	+	40.00	123.01	1.1e+03	542	1 BE750188 201387 MARC 4BOV Bos t	gb_est1:AA435355	-	39.00	123.90	975.79	339	1 AA435355 v049h12.s1 Knowles
gb_est2:BG664529	-	40.00	122.96	1.2e+03	545	1 BG664529 DRABPA09 Rat DRG Libra	gb_est1:BE770834	-	39.00	123.77	992.22	344	1 BE770834 CM2-TF0070-040700-2
gb_est2:BG909314	-	40.00	122.89	1.1e+03	549	1 BG909314 TaLR1175H09F TaLR1 Tr1	gb_est1:AW608121	-	39.00	123.71	998.80	346	1 AW608121 OV3-LT0048-270100-0
gb_est1:A1077307	+	40.00	122.77	1.1e+03	551	1 A1077307 OY99g10.x1 Soares_feta	gb_est1:AW608121	-	39.00	123.71	998.80	346	1 AW608121 OV3-LT0048-270100-0
gb_gss:AZ957674	-	40.00	122.75	1.1e+03	558	1 AZ957674 2M0224E08R Mouse 10kb	gb_est1:AO69744	-	39.00	123.66	1.0e+03	348	1 AO69744 AU069744 Rice panic
gb_est1:BI2770882	-	40.00	122.69	1.1e+03	562	1 BI2770882 606059G03.x2 606 - Ear	gb_est1:BB162825	-	39.00	123.59	1.0e+03	351	1 BB162825 BB162825 RIKEN full
gb_est2:BG905794	-	40.00	122.56	1.2e+03	567	1 BG905794 TaLR1142E03F TaLR1 Tr1	gb_gss:AO42443	-	39.00	123.39	1.0e+03	359	1 AO42443 CIT-HSP-23262F.TR C
gb_gss:AZ401890	-	40.00	122.55	1.2e+03	571	1 AZ401890 1M0168M19R Mouse 10kb	gb_est1:BI0347171	-	39.00	123.29	1.1e+03	363	1 BI0347171 tc04e01.s1 NCI CGAP
gb_est1:AI179963	-	40.00	122.51	1.2e+03	573	1 AI179963 EST223694 Normalized r	gb_est1:HS9004	-	39.00	123.12	1.1e+03	370	1 HS9004 yr40h01.r1 Soares fet
gb_est2:C87212	+	40.00	122.51	1.2e+03	573	1 C87212 C87212 Mouse fertilized	gb_est1:BI271686	-	39.00	122.97	1.1e+03	376	1 BI271686 q182803.x1 NCI CGAP
gb_est1:AA516602	-	40.00	122.48	1.2e+03	575	1 AA516602 vH85e04.r1 Knowles Sol	gb_est1:AV687719	-	39.00	122.86	1.1e+03	381	1 AV687719 AV687719 GK Homo s
gb_est2:C80683	-	40.00	122.34	1.2e+03	584	1 C80683 C80683 Mouse 3.5-dpc bla	gb_est1:AW62697	-	39.00	122.81	1.1e+03	383	1 AW62697 gal4C08.y1 Moss EST
gb_est1:BI275683	+	40.00	122.31	1.2e+03	586	1 BI275683 UI-R-CWO-bwm-e-01-0-UI	gb_est1:AW62697	-	39.00	122.76	1.1e+03	383	1 AW62697 gal4C08.y1 Moss EST
gb_est1:BE213409	+	40.00	122.15	1.2e+03	597	1 BE213409 EST0178 Triticum aesti	gb_est2:BI054971	-	39.00	122.67	1.1e+03	389	1 BI054971 CM3-GN0330-050201-7
gb_gss:BH027176	-	40.00	122.12	1.2e+03	599	1 BH027176 RPI2-24-247D24.TJ RPI	gb_gss:AO997535	-	39.00	122.63	1.1e+03	391	1 AO997535 RPI2-23-383H24.TV R
gb_est2:BG712591	-	40.00	121.99	1.2e+03	608	1 BG712591 RPI2-24-117G20.TJ RPI	gb_gss:BH025381	-	39.00	122.54	1.2e+03	395	1 BH025381 RPI2-24-335H24.TV R
gb_est1:BG12931	+	40.00	121.97	1.2e+03	609	1 BG12931 P911p_pk014.n2 Normali	gb_est1:AI044824	-	39.00	122.51	1.2e+03	396	1 AI044824 UI-R-C1-k1-e-03-0-U
gb_est2:BF282278	-	40.00	121.97	1.2e+03	609	1 BF282278 EST46957 Rat Gene Inc	gb_est1:AI044824	-	39.00	122.51	1.2e+03	396	1 AI044824 UI-R-C1-k1-e-03-0-U
gb_est2:BG905660	-	40.00	121.94	1.3e+03	611	1 BG905660 TaLR1141D17F TaLR1 Tr1	gb_est1:AL362566	-	39.00	122.42	1.2e+03	400	1 AL362566 AL362566 ICRP 522
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_gss:AZ449229	-	40.00	121.91	1.3e+03	621	1 AZ449229 IM0247D32F Mouse 10kb	gb_est2:BI010193	-	39.00	122.42	1.2e+03	400	1 BI010193 CM1-LT0042-100101-7
gb_g													



gb_estl:BE054797	39.00	114.77	3.1e+03	946	BE054797 GA_Ea0031G19f Gossyp	gb_est2:BF550641	-	38.00	119.36	1.7e+03	390	BF550641 UI-R-C0-hr-f-10-0-U
gb_est2:BF087162	39.00	114.77	3.1e+03	946	BI087162 6028504261 FTI_NH_MGC_10	gb_est2:T18864	+	38.00	119.07	1.8e+03	403	T18864 h06019t nestbis l Homo
gb_estl:A1535816	39.00	114.77	3.2e+03	952	AL1535816 LTI_NH_MGC_10	gb_estl:AA65081	+	38.00	119.01	1.8e+03	406	AA65081 UI-R-C0-gz-b-03-0-U
gb_estl:A1519809	39.00	114.57	3.2e+03	967	AL1519809 LTI_NH_MGC_10	gb_estl:AW77964	+	38.00	118.94	1.8e+03	409	AW77964 f4k4204.01 zebrafis
gb_est2:BF0107106	39.00	114.57	3.3e+03	990	BG107106 602291010f NIH_MGC_8	gb_estl:AF114097	+	38.00	118.90	1.9e+03	411	AF114097 AF114097 Homo sapie
gb_gss:CN503614	39.00	114.06	3.4e+03	1025	AL154785 602291010f NIH_MGC_8	gb_estl:AI580407	+	38.00	118.88	1.9e+03	412	AI580407 tm42h04.x1 NCI_CGAP
gb_est2:BF348255	39.00	113.83	3.6e+03	1052	BF348255 60202179f1 NCI_CGAP	gb_estl:AW405257	+	38.00	118.88	1.9e+03	412	AW405257 UI-HF-BK0-gas-f-02-
gb_est2:BF3530867	39.00	113.73	3.6e+03	1063	BF3530867 60207252f1 NCI_CGAP	gb_estl:AW478499	+	38.00	118.85	1.9e+03	413	AW478499 20541 MARC 1BOV Bos
gb_estl:BE040722	39.00	113.71	3.6e+03	1066	BE040722 OF05F06 OF Orviza satl	gb_estl:AW834878	+	38.00	118.79	1.9e+03	416	AW834878 RC6-LT0003-170100-Bos
gb_gss:CN503637	39.00	113.50	3.7e+03	1091	AL2330114 Tetraodon nigroviridi	gb_estl:BS071782	+	38.00	118.79	1.9e+03	416	BS071782 RC6-BT0518-131299-0
gb_est2:BF0338340	39.00	113.24	3.8e+03	1124	BG338340 602436120f1 NIH_MGC_4	gb_est2:R22238	+	38.00	118.77	1.9e+03	417	R22238 yh26a10.sl Soares pia
gb_htc:AK009601	39.00	112.39	4.3e+03	1236	AK009601 Mus musculus adult ma	gb_estl:AI1819667	+	38.00	118.68	1.9e+03	421	AI1819667 wJ59g02.x1 NCI_CGAP
gb_est2:BF623596	39.00	111.16	5.0e+03	1420	BG623596 602728986f1 NIH_MGC_1	gb_estl:BE099237	+	38.00	118.68	1.9e+03	421	BE099237 UI-R-BJ1-atz-h-10-0
gb_htc:AK016037	39.00	110.97	5.1e+03	1450	AK016037 Mus musculus adult ma	gb_est2:BF474013	+	38.00	118.64	1.9e+03	423	BF474013 WHE0840.A04.B0825 W
gb_htc:AK016425	39.00	102.50	1.5e+04	3760	AK016425 Mus musculus adult ma	gb_est2:BF524582	+	38.00	118.64	1.9e+03	423	BF524582 UI-R-AA0-ws-g-10-0-
gb_htc:BC009948	39.00	100.50	2.0e+04	4708	BC009948 Homo sapiens clone1	gb_estl:AV410098	+	38.00	118.56	1.9e+03	427	AV410098 AV410098 Lotus japo
gb_est2:BF6688489	38.50	124.72	878.43	257	BF644378 RC5-NN1164-131000-011	gb_estl:AW128224	+	38.00	118.47	2.0e+03	431	AW128224 UI-R-B01-aju-e-10-0
gb_estl:AW653600	38.50	119.52	1.7e+03	461	BG688489 336135 BARC 5BOV Bos t	gb_estl:AW490898	+	38.00	118.43	2.0e+03	433	AW490898 UI-M-BH3-att-d-02-0
gb_estl:BE258748	38.50	116.15	2.0e+03	538	AW653600 102365 MARC 1BOV Bos t	gb_est2:BF57679	+	38.00	118.43	2.0e+03	433	BF57679 yr27b11.sl Soares fet
gb_gss:AZ806213	38.00	124.78	870.71	212	BE258748 60113752f1 NIH_MGC_16	gb_est2:BF524968	+	38.00	118.41	2.0e+03	434	BF524968 UI-R-A0-af-c-06-0-U
gb_estl:AA0267541	38.00	130.22	433.44	115	AZ806213 2M0068C12f Mouse 10kb	gb_est2:N76851	+	38.00	118.39	2.0e+03	435	N76851 yv48h02.r1 Soares fet
gb_gss:AZ839001	38.00	128.60	533.61	138	AA0267541 mi06a02.r1 Soares mous	gb_gss:AO686867	+	38.00	118.37	2.0e+03	436	AO686867 nbx0073019r CUGI R
gb_estl:BE0339785	38.00	126.10	569.03	146	AZ839001 2M0134124R Mouse 10kb	gb_gss:AO706257	+	38.00	118.33	2.0e+03	438	AO706257 HS_3553.AL.E08.T7A
gb_estl:AV005499	38.00	125.57	786.91	184	BE0339785 RB039785 RIKEN full-le	gb_est2:HB1045	+	38.00	118.17	2.0e+03	446	HB1045 yu60d03.sl Soares fet
gb_estl:AI1050685	38.00	124.78	870.71	212	AV005499 AV005499 Mus musculus	gb_gss:AO7371725	+	38.00	118.13	2.0e+03	448	AO7371725 HS_3380.B2.G04.SP6E
gb_est2:BG723352	38.00	124.02	960.25	231	AI1050685 ov42c03.sl Soares test	gb_estl:AI1704112	+	38.00	118.07	2.1e+03	451	AI1704112 UI-R-AF0-yc-e-12-0-
gb_est2:BG772922	38.00	124.02	960.25	231	BG723352 602694007f1 NIH_MGC_97	gb_est2:BI7005179	+	38.00	118.07	2.1e+03	451	BI7005179 PM3-HN0076-020401-0
gb_estl:AV431391	38.00	123.76	993.51	238	BG772922 602721152f1 NIH_MGC_97	gb_est2:N90464	+	38.00	118.05	2.1e+03	452	N90464 zal7c01.r1 Soares fet
gb_estl:BA468954	38.00	123.21	1.1e+03	253	AV431391 AV431391 Porphyra yezo	gb_estl:BA099910	+	38.00	118.01	2.1e+03	454	BA099910 UI-R-BJ1-ati-e-10-0
gb_estl:AW341589	38.00	122.97	1.2e+03	260	BA468954 BA468954 RIKEN full-le	gb_estl:BF453879	+	38.00	118.01	2.1e+03	454	BF453879 maa72h09.y1 Soares
gb_est2:H33457	38.00	122.50	1.2e+03	274	AW341589 hdlf11.x1 Soares NFL	gb_estl:AL507996	+	38.00	117.99	2.1e+03	455	AL507996 AL507996 Hordeum vu
gb_estl:BF152845	38.00	122.09	1.3e+03	282	H33457 ESPI09475 Rat PC-12 cell	gb_est2:H20571	+	38.00	117.95	2.1e+03	457	H20571 ym47h08.sl Soares inf
gb_estl:BF152845	38.00	121.94	1.3e+03	287	BF152845 uz96f12.y1 NCI_CGAP	gb_estl:AI158054	+	38.00	117.93	2.1e+03	458	AI158054 ud26a04.r1 Soares t
gb_estl:AA482568	38.00	121.91	1.3e+03	293	BF152845 uz96f12.y1 NCI_CGAP	gb_estl:AW531334	+	38.00	117.90	2.1e+03	460	AW531334 UI-R-BU0-ane-c-05-0
gb_estl:BA482568	38.00	121.88	1.3e+03	294	AA482568 zt34e08.sl Soares ovan	gb_est2:BF931618	+	38.00	117.82	2.1e+03	464	BF931618 IL2-NT0203-141200-3
gb_estl:BB0514566	38.00	121.85	1.3e+03	295	BB514566 BB514566 RIKEN full-le	gb_estl:AA549217	+	38.00	117.80	2.1e+03	466	AA549217 VK74F02.sl Knowles
gb_estl:BB034757	38.00	121.79	1.3e+03	297	BB34757 BB234757 RIKEN full-le	gb_est2:BF32681	+	38.00	117.78	2.1e+03	466	BF32681 nael1a03.x1 NCI_CGA
gb_estl:BB006104	38.00	121.64	1.3e+03	302	BB006104 BB006104 RIKEN full-le	gb_gss:AO821257	+	38.00	117.76	2.1e+03	467	AO821257 UI-R-CS0-btq-h-03-0
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_gss:AO821719	+	38.00	117.76	2.1e+03	467	AO821719 HS_5431.A2.D03.SP6E
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_gss:AO138855	+	38.00	117.74	2.1e+03	468	AO138855 HS_3075.A2.F03.MF C
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203	38.00	121.52	1.4e+03	306	BF120203 UI-R-YO-vg-g-07-0-U	gb_estl:AV641738	+	38.00	117.67	2.2e+03	472	AV641738 AV641738 Chlamydomo
gb_estl:BF120203												



gb_gss:BH056025	38.00	116.67	2.5e+03	528	1	BH056025	RPCI-24-365A16.TJ.RPCI	gb_est2:BE898980	38.00	112.70	4.1e+03	825	1	BE898980	601512241F1.NIH.MGC
gb_gss:Q727888	38.00	116.59	2.5e+03	533	1	Q727888	MS_2258.A2.B02.MR.CIT	gb_gss:AQ573945	38.00	112.63	4.1e+03	832	1	AQ573945	nbxb0083601F1.CUGI.R
gb_gss:BF660799	38.00	116.54	2.5e+03	536	1	BF660799	msa72h09.x1.Soaress.mou	gb_est2:W08409	38.00	112.59	4.2e+03	835	1	W08409	mb49a10.x1.Soaress.mou
gb_gss:AZ382398	38.00	116.54	2.5e+03	536	1	AZ382398	1M0139L08R.Mouse.10kb	gb_est2:BG537584	38.00	112.56	4.2e+03	838	1	BG537584	602565916F1.NIH.MGC
gb_gss:AA967290	38.00	116.50	2.5e+03	538	1	AA967290	vz39h009.r1.Soaress.chym	gb_est1:AL3661129	38.00	112.52	4.2e+03	842	1	AL3661129	LTI.NFL010
gb_gss:C78770	38.00	116.50	2.5e+03	538	1	C78770	C78770.Mouse.3.5-dpc.bla	gb_est1:AW3821218	38.00	112.46	4.2e+03	848	1	AW3821218	HVSM69000E23f.Hord
gb_est2:C78770	38.00	116.49	2.5e+03	539	1	CG040075	NXSI_106.B09.F.NXSI.(N	gb_gss:AQ573615	38.00	112.42	4.3e+03	852	1	AQ573615	nbxb0084F19r.CUGI.R
gb_gss:AQ047221	38.00	116.49	2.5e+03	539	1	AQ047221	RPCI11-35P16.TJ.RPCI.(N	gb_hcc:AK020369	38.00	112.38	4.3e+03	855	1	AK020369	Mus.musculus.adult
gb_gss:AZ487901	38.00	116.47	2.5e+03	539	1	AZ487901	1M0317B20R.Mouse.10kb	gb_gss:AQ245374	38.00	112.37	4.3e+03	855	1	AQ245374	nbxb004907r.CUGI.R
gb_gss:AQ041210	38.00	116.47	2.5e+03	540	1	AQ041210	RPCI11-20E2.TKBF.RPCI-	gb_gss:AQ245374	38.00	112.37	4.3e+03	856	1	AQ245374	HS_2060.B1.C01.T7.C
gb_est1:AW622560	38.00	116.39	2.6e+03	545	1	AW622560	EST313360.tomato.root	gb_est1:AL346799	38.00	112.35	4.3e+03	858	1	AL346799	AL546799.LTI.NFL006
gb_gss:AZ375161	38.00	116.32	2.6e+03	548	1	AZ375161	1M0128G19F.Mouse.10kb	gb_est2:BG529094	38.00	112.32	4.3e+03	861	1	BG529094	602780803F1.NCI.CGA
gb_gss:AZ621879	38.00	116.32	2.6e+03	549	1	AZ621879	1M045016F.Mouse.10kb	gb_gss:CN500A08	38.00	112.30	4.3e+03	863	1	CN500A08	AL034974.Drosophila.melanoga
gb_est1:BE755889	38.00	116.26	2.6e+03	553	1	BE755889	209768.MARC.2BOV.Bos.t	gb_est1:AL525921	38.00	112.28	4.4e+03	865	1	AL525921	LTI.NFL003
gb_est1:AL040324	38.00	116.10	2.7e+03	563	1	AL040324	DFE2743.NC1113.r1.434	gb_est1:BE569621	38.00	112.07	4.4e+03	886	1	BE569621	601328268F1.NCI.CGA
gb_gss:AZ994238	38.00	116.05	2.7e+03	566	1	AW944238	SD02424.Sprime.SD.Dros	gb_est2:BI102559	38.00	112.07	4.4e+03	886	1	BI102559	602888119F1.NCI.CGA
gb_est2:BI346612	38.00	116.02	2.7e+03	568	1	BI346612	375864.MARC.2PIC.Sus.s	gb_gss:AQ245374	38.00	112.07	4.4e+03	886	1	AQ245374	HS_2537.A2.B01.T7A
gb_est1:BI195842	38.00	115.97	2.7e+03	571	1	AI995842	70154866.A.A.thaliana,	gb_gss:AQ747438	38.00	111.85	4.6e+03	905	1	AQ747438	HS_5537.A2.B01.T7A
gb_gss:AV607797	38.00	115.94	2.7e+03	573	1	AV607797	AV607797.Bos.taurus.ki	gb_est2:BE971354	38.00	111.81	4.6e+03	912	1	BE971354	601651521F1.NIH.MGC
gb_est1:AA851938	38.00	115.85	2.7e+03	579	1	AA851938	EST194707.Normalized.r	gb_gss:CN50610J	38.00	111.68	4.7e+03	925	1	CN50610J	AL399689.T7.end.of.clone.ASO
gb_est2:BG52489	38.00	115.77	2.8e+03	584	1	BG52489	RHI22_13.F11.b1.A003.R	gb_est1:BE270919	38.00	111.64	4.7e+03	930	1	BE270919	600943796F1.NIH.MGC
gb_est1:AA034752	38.00	115.74	2.8e+03	586	1	AA034752	m144b11.r1.Soaress.mous	gb_gss:CN501EVK	38.00	111.56	4.7e+03	938	1	CN501EVK	AL141057.Anopeles.gambiae.G
gb_gss:AZ971694	38.00	115.70	2.8e+03	589	1	AZ971694	2M0245J13F.Mouse.10kb	gb_gss:CN500GGG	38.00	111.29	4.9e+03	967	1	CN500GGG	AL397664.T7.end.of.clone.ASO
gb_gss:AZ993364	38.00	115.67	2.8e+03	591	1	AZ993364	2M027BD20F.Mouse.10kb	gb_gss:CN503LI4	38.00	111.19	5.0e+03	989	1	CN503LI4	AL249493.Tetraodon.nigroviri
gb_gss:AZ993364	38.00	115.65	2.8e+03	591	1	AZ993364	2M027BD20F.Mouse.10kb	gb_est1:BE745770	38.00	111.09	5.0e+03	989	1	BE745770	601579957F1.NIH.MGC
gb_est1:AU006094	38.00	115.55	2.8e+03	599	1	AU006094	AU006094.Bombyx.mori.F	gb_gss:CN503UBF	38.00	110.98	5.1e+03	1001	1	CN503UBF	AL260916.Tetraodon.nigrovir
gb_est2:BE910827	38.00	115.52	2.8e+03	599	1	BE910827	601661889F1.NCI.CGAP.M	gb_gss:CN505RL8	38.00	110.77	5.1e+03	1001	1	CN505RL8	AL350693.Tetraodon.nigrovir
gb_est1:BE910827	38.00	115.52	2.8e+03	599	1	BE910827	601661889F1.NCI.CGAP.M	gb_est2:BF624742	38.00	110.63	5.3e+03	1041	1	BF624742	HVSM6A0017J22f.Hor
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5e+03	1070	1	CN500G8V	AL408103.602918905F1.NCI.CG
gb_gss:AZ974011	38.00	115.52	2.9e+03	601	1	AW011551	SRT2B07.Pine.TF.CI	gb_gss:CN500G8V	38.00	110.39	5.5				

gb_est2:BF925624	-	37.00	118.80	1.9e+03	287	BF925624	CM2-NT0192-291100-578-	gb_est2:D26782	+	37.00	115.40	2.9e+03	421	D26782	CELK002DGR	Yuji Kohar	
gb_gss:AB328570	+	37.00	118.80	1.9e+03	287	AB328570	IM0052H24F	Mouse 10kb	-	37.00	115.33	2.9e+03	424	AA147516	z150f02.r1	Soares_p	
gb_est1:BB339475	+	37.00	118.77	1.9e+03	288	B339475	BB339475	R1KEN full-le	-	37.00	115.33	2.9e+03	424	BE446737	WHE1140	H11_P22S_W	
gb_est1:BB756031	+	37.00	118.74	1.9e+03	289	AW756031	s1112g1.v1	Gm-cl036	GL	-	37.00	115.27	2.9e+03	427	BF471720	UI-M-BH3-avw-c-08-0	
gb_est1:BB295953	+	37.00	118.65	1.9e+03	292	B295953	BB295953	R1KEN full-le	-	37.00	115.27	2.9e+03	427	BF651413	274196	MARC 3BOV	
gb_gss:AA884807	-	37.00	118.65	1.9e+03	292	AA884807	RPCI-23-191K12	TV RPCI	-	37.00	115.25	3.0e+03	428	BE523343	M35H10STM	Arabidops	
gb_est2:BF337996	+	37.00	118.62	1.9e+03	293	T93796	UI-R-CAL1	Stratagene	lu	-	37.00	115.21	3.0e+03	428	T62237	yb97h11.s1	Stratagene
gb_est2:BF387101	+	37.00	118.56	1.9e+03	295	BF387101	UI-R-CAL1	Stratagene	lu	-	37.00	115.21	3.0e+03	430	BE186864	946013R10	X1 946 -
gb_est1:AA467249	+	37.00	118.44	2.0e+03	298	AA467249	VF06a07.r1	Knowles	Sol	-	37.00	115.19	3.0e+03	430	BG989934	PM2-HT1171-2F	Mouse 10
gb_est2:BG014250	+	37.00	118.41	2.0e+03	300	BG014250	CM3-GN0296-211200-574-		-	37.00	115.15	3.0e+03	431	BE481504	166578	BARC 5BOV	
gb_est2:BG135743	+	37.00	118.29	2.0e+03	304	B3135743	sab84049.y1	Gm-cl032	G	-	37.00	115.13	3.0e+03	433	BE481504	166578	BARC 5BOV
gb_est1:BB259431	+	37.00	118.23	2.0e+03	306	B259431	BB259431	R1KEN full-le	-	37.00	115.13	3.0e+03	434	AI183413	au42d05.y1	Schneide	
gb_est1:BB541515	+	37.00	118.21	2.0e+03	307	BB541515	BB541515	R1KEN full-le	-	37.00	115.09	3.0e+03	436	AZ383647	RPCI-23-125F23	TV R	
gb_gss:TA48080P	+	37.00	118.21	2.0e+03	307	TA48080P	T. brucei	sheared	gend	-	37.00	115.09	3.0e+03	436	D33608	CELK033EAR	Yuji Kohar
gb_est1:BB241183	+	37.00	118.12	2.0e+03	310	BB241183	BB241183	R1KEN full-le	-	37.00	115.09	3.0e+03	436	D34120	CELK041CAR	Yuji Kohar	
gb_est1:BB361156	+	37.00	118.12	2.0e+03	310	BB361156	BB361156	R1KEN full-le	-	37.00	115.07	3.0e+03	437	BF469901	UI-M-BH3-atw-h-12-0		
gb_gss:AA050958	-	37.00	118.03	2.1e+03	313	AA050958	GSSRC11809	Trypanosoma		-	37.00	115.07	3.0e+03	437	BF808552	CM1-CI0092-141100-5	
gb_est1:AI381755	+	37.00	117.78	2.1e+03	322	AI381755	te40905.x1	Soares_NHMK		-	37.00	115.05	3.0e+03	438	BF606706	273820	MARC 3BOV
gb_est2:BF912630	+	37.00	117.75	2.1e+03	322	BF912630	MR3-UT0050-211100-002-		-	37.00	115.03	3.0e+03	439	D26780	CELK011GZR	Yuji Kohar	
gb_est2:BF912629	+	37.00	117.73	2.2e+03	324	BF912629	MR3-UT0050-211100-002-		-	37.00	115.01	3.1e+03	440	AA493914	nh07d09.s1	NCL_CGAP	
gb_est2:BI060450	+	37.00	117.70	2.2e+03	325	BI060450	MR3-UT0050-220101-001-		-	37.00	115.01	3.1e+03	440	BE945955	UI-M-BH3-avw-c-07-0		
gb_gss:AA886942	+	37.00	117.59	2.2e+03	329	AA886942	RPCI-23-195G16	TJ RPCI	-	37.00	114.99	3.1e+03	441	AW312777	5072	MARC 2BOV	
gb_est2:F00349	+	37.00	117.43	2.2e+03	335	F00349	HSB10G062	STRATAGENE	Hum	-	37.00	114.94	3.1e+03	443	AI369164	qy740R8	X1 NCL_CGAP
gb_est2:BF606705	+	37.00	117.43	2.2e+03	335	BF606705	273819	MARC 3BOV	Bos t	-	37.00	114.88	3.1e+03	443	B28288	T9D207R8	TAMU Arabidops
gb_est1:BF771932	+	37.00	117.30	2.3e+03	340	BF771932	IL5-IT0029-291100-259-		-	37.00	114.85	3.1e+03	446	AU166505	AU166505	Rice callu	
gb_est1:AW084942	-	37.00	117.25	2.3e+03	342	AW084942	xc65f02.x1	NCL_CGAP	Es	-	37.00	114.85	3.1e+03	448	AA027867	zk05b03.r1	Soares_p
gb_est1:BE525030	+	37.00	117.22	2.3e+03	343	BE525030	M57G1STM	Arabidopsis	d	-	37.00	114.85	3.1e+03	448	AZ164697	SP_0076	BL_F07_SP6E
gb_est2:T00280	+	37.00	117.22	2.3e+03	343	T00280	wEST01001	Early embryo,		-	37.00	114.85	3.1e+03	448	AQ544840	CITBI-EL-2612K4	FR
gb_est1:AW312775	+	37.00	117.19	2.3e+03	344	AW312775	5069	MARC 2BOV	Bos tau	-	37.00	114.83	3.1e+03	449	BF044495	BP2500301A01f1	Soar
gb_est2:D34104	+	37.00	117.14	2.3e+03	346	D34104	CELK041A6R	Yuji Kohara	u	-	37.00	114.83	3.1e+03	449	BF957382	RC4-N060231-241100-	
gb_est1:AI31859	+	37.00	117.12	2.3e+03	347	AI31859	C91859	Rice panicle	short	-	37.00	114.83	3.1e+03	449	AS824202	HS-3060A2	A03_MR C
gb_est2:BF425560	+	37.00	117.07	2.4e+03	349	BF425560	su44c04.y1	Gm-cl068	GL	-	37.00	114.81	3.1e+03	450	AW677350	DGL-5_H06_g1_A002	
gb_est2:BF109290	+	37.00	116.94	2.4e+03	354	BF109290	7151d12.x1	Soares_NSF		-	37.00	114.81	3.1e+03	450	AZ239868	RPCI-23-71G18	TV RP
gb_est2:BG511006	+	37.00	116.86	2.4e+03	357	BG511006	sac81d03.y1	Gm-cl072	G	-	37.00	114.79	3.1e+03	451	BG649255	EM1_79_E05_g1_A002	
gb_est1:AW296961	+	37.00	116.84	2.4e+03	358	AW296961	UI-H-BW0-aJc-rf-10-0-UI		-	37.00	114.75	3.2e+03	453	AW010366	ST05H05	Pine Triplee	
gb_est2:D32780	+	37.00	116.77	2.4e+03	361	D32780	CELK02108R	Yuji Kohara	u	-	37.00	114.71	3.2e+03	455	AA018781	ze51c10.r1	Soares_r
gb_est1:AW233115	+	37.00	116.69	2.5e+03	363	D33315	CELK0229B3R	Yuji Kohara	u	-	37.00	114.71	3.2e+03	455	AW797717	VMJ31c12.r1	Soares_m
gb_gss:BB7459	+	37.00	116.64	2.5e+03	366	B27459	T8L10TR	TAMU Arabidopsis		-	37.00	114.69	3.2e+03	456	AZ048075	LMJ2e11.x1	NCL_CGAP
gb_est1:AW91167	+	37.00	116.62	2.5e+03	367	AW91167	UI-M-BH3-atn-d-07-0-UI		-	37.00	114.67	3.2e+03	457	AI497585	tm66f06.x1	NCL_CGAP	
gb_est2:R31381	+	37.00	116.62	2.5e+03	367	R31381	YH75605.r1	Soares	placem	-	37.00	114.65	3.2e+03	457	AQ222504	RPCI-23-291K4	TV RP
gb_est1:D32760	+	37.00	116.52	2.5e+03	371	D32760	CELK02508R	Yuji Kohara	u	-	37.00	114.65	3.2e+03	458	BF566999	UI-R-BU0P-agb-g-06-	
gb_est1:AI616087	+	37.00	116.47	2.5e+03	373	AI616087	vr54h04.x1	Knowles	Sol	-	37.00	114.61	3.2e+03	460	AI189015	qd27h03.x1	Soares_p
gb_gss:AA284926	+	37.00	116.40	2.6e+03	376	AA284926	2M0150N01R	Mouse 10kb		-	37.00	114.61	3.2e+03	460	AI278221	q18e07.x1	Soares_N
gb_est1:AA982624	+	37.00	116.29	2.6e+03	381	AA982624	uh11c03.r1	Soares	mous	-	37.00	114.57	3.2e+03	462	AI189551	qd28a04.x1	Soares_p
gb_est1:AW036015	+	37.00	116.22	2.6e+03	384	AW036015	EST283099	tomato	callu	-	37.00	114.55	3.2e+03	463	D33654	CELK034B3R	Yuji Kohar
gb_est1:AI901428	+	37.00	116.17	2.6e+03	386	AI901428	61800A11.x1	618 - inh		-	37.00	114.53	3.2e+03	464	AW58598	HS_2034_B1_H08_MR C	
gb_est1:AW221127	+	37.00	116.12	2.6e+03	388	AW221127	EST297596	tomato	fruit	-	37.00	114.49	3.3e+03	466	AW034869	EST279098	tomato ca
gb_est2:BI048334	+	37.00	116.01	2.7e+03	392	BI048334	PM2-ST0303-220401-013-		-	37.00	114.48	3.3e+03	467	AZ55962	LM0095C19R	Mouse 10	
gb_est1:D33472	+	37.00	116.01	2.7e+03	393	D33472	CELK031E3R	Yuji Kohara	u	-	37.00	114.48	3.3e+03	467	AZ557026	RPCI-23-205H15	TV R
gb_gss:AA722103	-	37.00	116.01	2.7e+03	393	AZ722103	RPCI-24-93G17	TV RPCI	-	37.00	114.44	3.3e+03	469	AZ794349	2M0048P05F	Mouse 10	
gb_est2:D24436	+	37.00	115.94	2.7e+03	396	D24436	RICR1888A	Rice root	Oryz	-	37.00	114.36	3.3e+03	473	BF588533	7102h08.x1	NCL_CGAP
gb_est1:D323018	+	37.00	115.92	2.7e+03	397	D323018	CELK025A6R	Yuji Kohara	u	-	37.00	114.31	3.3e+03	476	BE344722	946028B09.y1	946 -
gb_est1:AA493942	+	37.00	115.87	2.7e+03	398	AA493942	UI-M-BH3-auh-h-08-0-UI		-	37.00	114.29	3.3e+03	477	BF021748	uy57f02.y1	NCL_CGAP	
gb_gss:AA150216	+	37.00	115.88	2.7e+03	399	AA150216	SP_0049_B2_F11-SP6E	St		-	37.00	114.29	3.3e+03	477	BF116964	uz04a05.y1	NCL_CGAP
gb_gss:BB32110	+	37.00	115.85	2.7e+03	400	B32110	HS-1014-rl-H01-MR	abi	Cl	-	37.00	114.27	3.4e+03	478	AW027247	yv66c04.x1	Soares_t
gb_est2:AA924227	+	37.00	115.83	2.7e+03	401	AZ924227	4906-ic27p08.s1	Saccha		-	37.00	114.27	3.4e+03	478	AW027247	yv66c04.x1	Soares_t
gb_est2:D34398	+	37.00	115.81	2.8e+03	402	D34398	CELK04503R	Yuji Kohara	u	-	37.00	114.27	3.4e+03	478	AZ251480	RPCI-23-466B23	TUB
gb_est1:BF146085	+	37.00	115.81	2.8e+03	402	BF146085	WHE1837_A05_A09Z5	Seca		-	37.00	114.27	3.4e+03	478	BH12105	RPCI-24-402023	TJ R
gb_est1:AA1887059	+	37.00	115.79	2.8e+03	403	AA1887059	w196C02.x1	NCL_CGAP	Br	-	37.00	114.25	3.4e+03	479	AA96934	UI-R-CO-hj-g-09-0-U	
gb_est2:D33412	+	37.00	115.76	2.8e+03	404	D33412	CELK044B4R	Yuji Kohara	u	-	37.00	114.25	3.4e+03	479	AA96934	UI-R-CO-hj-g-09-0-U	
gb_est2:D36778	+	37.00	115.74	2.8e+03	405	D36778	CELK01685R	Yuji Kohara	u	-	37.00	114.21	3.4e+03	481	AI341757	qq91c10.x1	Soares_t
gb_est2:BF324709	+	37.00	115.74	2.8e+03	405	BF324709	sui14d04.y1	Gm-cl066	GL	-	37.00	114.21	3.4e+03	481	AW373246	RC5-BT0508-031299-0	
gb_est1:AW14580	+	37.00	115.72	2.8e+03	406	AW14580	fj05h10.y1	Zebrafish	a	-	37.00	114.21	3.4e+03	481	BI073586	kt34e05.y1	Strongyl
gb_est1:AW216186	+	37.00	115.68	2.8e+03	408	AW216186	687043A01.y1	687 - Ear		-	37.00	114.16	3.4e+03	484	BE502586	hz18b09.x1	NCL_CGAP
gb_est1:AI380143	+	37.00	115.63	2.8e+03	410	AI380143	tf93a05.x1	NCL_CGAP	Br	-	37.00	114.16	3.4e+03	484	AZ293959	np213-163018	TJ R
gb_est1:BE523805	+	37.00	115.63	2.8e+03	410	BE523805	M41H12STM	Arabidopsis		-	37.00	114.14	3.4e+03	485	AA610527	rp91f07.s1	NCL_CGAP
gb_est1:AA584854	+	37.00	115.59	2.8e+03	412	AA584854	no09h11.s1	NCL_CGAP	Ph	-	37.00	114.14	3.				



gb\_est2:BF558068 - 37.00 114.05 3.5e+03 490 1 BF558068 UI-R-CO-hj-g-09-0-UI.1  
 gb\_est1:AF083303 + 37.00 114.03 3.5e+03 491 1 AF083303 AF083303 Drosophila me  
 gb\_est1:AI94302 - 37.00 114.03 3.5e+03 491 1 AI94302 QY98e10.x1 NCI\_CGAP\_B1  
 gb\_gss:BH085545 - 37.00 114.03 3.5e+03 491 1 BH085545 RPI-24-310A2.TV RPI-1  
 gb\_gss:AQ674982 + 37.00 114.01 3.5e+03 492 1 AQ674982 HS\_2161.AL\_HOI.T7C CIT-1  
 gb\_est1:AW586740 + 37.00 113.98 3.5e+03 494 1 AW586740 EST318363 MHAM Medica  
 gb\_est2:BF497115 + 37.00 113.98 3.5e+03 494 1 BF497115 AT111189.5Prime AT Dros  
 gb\_gss:AZ344778 + 37.00 113.98 3.5e+03 494 1 AZ344778 LM0079E02F Mouse 10kb  
 gb\_gss:AQ037915 + 37.00 113.96 3.5e+03 495 1 AQ037915 CIT-HSP-2325G5.TV CIT-  
 gb\_est1:AA394318 + 37.00 113.94 3.5e+03 496 1 AA394318 25915 Lambda-PRL2 Arad

seq\_name: gb\_est1:BE659955

seq\_documentation\_block: 711 bp mRNA EST 06-SEP-2000  
 LOCUS BE659955  
 DEFINITION 1077 GmaxSC Glycine max cDNA, mRNA sequence.  
 ACCESSION BE659955  
 VERSION BE659955.1 GI:9985949  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE 1 (bases 1 to 711)  
 AUTHORS Harris, N., Chapman, B.P. and Gijzen, M.  
 TITLE Gene expression in developing soybean seed coats  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Gijzen M  
 Agriculture and Agri-Food Canada  
 1391 Sandford Street, London, Ontario, Canada N5V 4T3  
 Tel: 519 457 1470  
 Fax: 519 457 3997  
 Email: gijzenm@agr.ca.  
 Location/Qualifiers  
 1..711

FEATURES  
 source  
 /organism="Glycine max"  
 /cultivar="Harosoy 63"  
 /db\_xref="taxon:3847"  
 /clone\_lib="GmaxSC"  
 /tissue\_type="Seed coats"  
 /lab\_host="E. coli strain XL0LR"  
 /note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI; This cDNA library was constructed from polyA+ enriched mRNA from green seed coats in mid to late developmental stage, average fresh weight 250 mg per seed. Traces of pod and embryo tissue also present. Complementary DNA was synthesized from mRNA using an XhoI-poly(dT) linker-primer. EcoRI adapters and the products were digested with XhoI for directional cloning into Lambda ZAP Express vector. This lambda library was amplified once using E. coli host strain XL1 Blue MRF'. Inserts were then subcloned by mass excision using ExAssist helper phage for conversion into phagemid vector pBK-CMV in E. coli host strain XL0LR."

BASE COUNT 186 a 207 c 179 g 134 t 5 others  
 ORIGIN  
 alignment\_scores:  
 Quality: 81.00 Length: 17  
 Ratio: 4.765 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-485-529-104 x BE659955 ..  
 Align seg 1/1 to: BE659955 from: 1 to: 711

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||||||||||||||||||||||||||||||||||||||||||||||||

174 GACGAGTGTGTCGGCGGTGGGTACAAGGTTCTGCTCCGACATGGC 223  
 17 a 17  
 224 T 224

seq\_name: gb\_est2:BF268018

seq\_documentation\_block: 844 bp mRNA EST 09-MAR-2001  
 LOCUS BF268018  
 DEFINITION HV\_CEA0019J17f Hordeum vulgare seedling green leaf EST library  
 HVCNDA0004 (Erysiphe infected & control) Hordeum vulgare cDNA clone  
 HV\_CEA0019J17f, mRNA sequence.

ACCESSION BF268018  
 VERSION BF268018.2 GI:13263734  
 KEYWORDS EST.  
 SOURCE Hordeum vulgare  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
 ; Triticeae; Hordeum.  
 REFERENCE 1 (bases 1 to 844)  
 AUTHORS Wing, R., Close, R.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu  
 , Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo  
 , T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and  
 Wood, T.  
 TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics  
 JOURNAL Unpublished (2000)  
 COMMENT On Nov 17, 2000 this sequence version replaced gi:11199013.  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: AATTAACCCCTCACTAAAGGG  
 High quality sequence stop: 587.

FEATURES  
 source  
 1..844  
 /organism="Hordeum vulgare"  
 /cultivar="CI16155 (Mlai3)"  
 /db\_xref="taxon:4513"  
 /clone\_lib="HV\_CEA0019J17f"  
 /tissue\_type="seedling green leaf"  
 /lab\_host="TJC121"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI; For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley/ To order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT 150 a 290 c 276 g 128 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 81.00 Length: 17  
 Ratio: 4.765 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-485-529-104 x BF268018 ..  
 Align seg 1/1 to: BF268018 from: 1 to: 844

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||||||||||||||||||||||||||||||||||||||||||||||||  
 260 GACGAGTGTGTCGGCGGTGGGTACAAGTCCGCGCTCCGACATGGC 309  
 17 a 17  
 |

310 G 310

seq\_name: gb\_est2:D39460

seq\_documentation\_block: 263 bp mRNA EST 11-NOV-1994  
LOCUS D39460 RICE SHOOT Oryza sativa cDNA, mRNA sequence.  
DEFINITION RICS0803A Rice shoot Oryza sativa cDNA, mRNA sequence.  
ACCESSION D39460  
VERSION D39460.1 GI:568611  
KEYWORDS EST.  
SOURCE Oryza sativa.  
ORGANISM Lotus japonicus.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 263)  
AUTHORS Sasaki,T., Miyao,A. and Yamamoto,K.  
TITLE Rice cDNA from callus 1995  
JOURNAL Unpublished (1995)  
COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan

Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/.  
FEATURES Location/Qualifiers  
source 1..263  
/organism="Oryza sativa"  
/strain="Nipponbare, sub\_species Japonica"  
/db\_xref="taxon:4530"  
/clone\_lib="Rice shoot"  
/note="Etiolated shoot (8 days old)"

BASE COUNT 46 a 64 c 116 g 35 t 2 others  
ORIGIN

alignment\_scores: Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block: US-09-485-529-104 x D39460 ..  
Align seg 1/1 to: D39460 from: 1 to: 263

1 AspGluLeuAlaLaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
97 GACGAGCTGTCGGCGCGCTGCGGTACAAAGTGGCTCGCGCATGGC 146

17 a 17  
147 C 147

seq\_name: gb\_est1:AV410222

seq\_documentation\_block: 355 bp mRNA EST 23-MAY-2000  
LOCUS AV410222 Lotus japonicus young plants (two-week old) Lotus  
DEFINITION AV410222 Lotus japonicus young plants (two-week old) Lotus  
japonicus cDNA clone MWL069f02\_r 5', mRNA sequence.

ACCESSION AV410222  
VERSION AV410222.1 GI:7723076  
KEYWORDS EST.  
SOURCE Lotus japonicus.  
ORGANISM Lotus japonicus.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;  
Lotus.

REFERENCE 1 (bases 1 to 355)  
AUTHORS Asanizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
TITLE Generation of 7137 non-redundant expressed sequence tags from a

JOURNAL Legume, Lotus japonicus  
MEDLINE DNA Res. 7 (2), 127-130 (2000)  
COMMENT 20277479

Contact: Yasukazu Nakamura  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES Location/Qualifiers  
source 1..355  
/organism="Lotus japonicus"  
/db\_xref="taxon:34305"  
/clone="MWL069f02\_r"  
/clone\_lib="Lotus japonicus young plants (two-week old)"  
/dev\_stage="Lotus japonicus young plants (two-week old)"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI; isolate=Miyakojima MG-20"

BASE COUNT 94 a 70 c 101 g 90 t  
ORIGIN

alignment\_scores: Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block: US-09-485-529-104 x AV410222 ..  
Align seg 1/1 to: AV410222 from: 1 to: 355

1 AspGluLeuAlaLaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
241 GATGAGCTTCTGGCGCTTAGGTACAAAGTGGCTTCCTCTGACATGGC 290

17 a 17  
291 T 291

seq\_name: gb\_est1:BE321891

seq\_documentation\_block: 372 bp mRNA EST 21-DEC-2000  
LOCUS BE321891 NF045B07IN1F1058 Insect herbivory Medicago truncatula cDNA clone  
DEFINITION NF045B07IN 5', mRNA sequence.

ACCESSION BE321891  
VERSION BE321891.2 GI:11964088  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 372)  
AUTHORS Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores  
H.R., Inman,J.T., Weller,J.W. and May,G.D.  
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula insect herbivory library  
JOURNAL Unpublished (2000)  
COMMENT On Jul 14, 2000 this sequence version replaced g1:9195668.

Contact: Korth K  
Dept. of Plant Pathology  
University of Arkansas  
217 Plant Science Building, Fayetteville, AR 72701, USA  
Tel: 501 575 5191  
Fax: 501 575 7601  
Email: kkorthe@comp.uark.edu  
Medicago Genome Initiative accession: MGI:S:23522  
Insert length: 798 Std Error: 0.00  
Plate: 045 row: B column: 07  
Seq primer: TCACACGAGGAACAGCATGAC.  
FEATURES Location/Qualifiers

```

source
1. .372
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF045B07IN"
/tissue_type="local and systemic leaves"
/dev_stage="mature"
/note="vector: Lambda Zap; Library was produced from fully
expanded M. truncatula leaves of plants fed upon by
Spodoptera exigua (beet armyworm) for 24 hours. Systemic
(undamaged leaves from injured plants) and wounded leaves
were harvested and pooled."
153 a 66 c 72 g 81 t
BASE COUNT 153 a 66 c 72 g 81 t
ORIGIN

alignment_scores:
Quality: 78.00 Length: 17
Ratio: 4.588 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-485-529-104 x BE321891 ..
Align seg 1/1 to: BE321891 from: 1 to: 372
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
247 GACGAGTCTTACGACCATAGTTACAGGTCGCTCTCCGACATGCG 296
17 a 17
297 T 297

seq_name: gb_est2:C27475
seq_documentation_block:
LOCUS C27475 388 bp mRNA EST 06-AUG-1997
DEFINITION C27475 Rice callus cDNA Oryza sativa cDNA clone C51976_1A, mRNA
sequence.
ACCESSION C27475
VERSION C27475.1 GI:2311320
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 388)
AUTHORS Yamamoto,K. and Sasaki,T.
TITLE Rice cDNA from callus 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/.
FEATURES
Location/Qualifiers
1. .388
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="C51976_1A"
/clone_lib="Rice callus cDNA"
/tissue_type="callus"
/dev_stage="callus"
BASE COUNT 83 a 102 c 134 g 65 t 4 others
ORIGIN

alignment_scores:

```

```

Quality: 78.00 Length: 17
Ratio: 4.588 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-485-529-104 x C27475 ..
Align seg 1/1 to: C27475 from: 1 to: 388
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
271 GACGAGCTGCTCGCGCGCTCGGTACAAGTCCGCTCGTCGACATGCG 320
17 a 17
321 C 321

seq_name: gb_est1:AL369357
seq_documentation_block:
LOCUS AL369357 399 bp mRNA EST 03-AUG-2000
DEFINITION MEB3A30E09F1 MtBA Medicago truncatula cDNA clone MtBA30E09 T3, mRNA
sequence.
ACCESSION AL369357
VERSION AL369357.1 GI:9669110
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolaeae;
Medicago.
REFERENCE 1 (bases 1 to 399)
AUTHORS Journet,E.P., Crespeau,H., van-Tuinen,D., Gouzy,J., Jaillon,O.,
Niebel,A., Carreau,V., Chatagnier,O., Kahn,D., Gianinazzi-Pearson,
V. and Gamas,P.
TITLE Medicago truncatula ESTs from nitrogen-starved roots
JOURNAL Unpublished (2000)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
Biologie Moleculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
Mt-est@toulouse.inra.fr Website :
http://sequence.toulouse.inra.fr/Mtruncatula.html).
FEATURES
Location/Qualifiers
1. .399
/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MEB3A30E09"
/clone_lib="MtBA"
/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="vector: pBluescript psk; Site.1: EcoRI; Site.2:
XhoI; Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-zapXR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExAssit helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."
BASE COUNT 166 a 74 c 75 g 84 t
ORIGIN

alignment_scores:

```

Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AL369357 ..

Align seg 1/1 to: AL369357 from: 1 to: 399

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
285 GACGAGTGTAGCAGCATAGGTACAAAGGTCGCTTCGACATGGC 334

17 a 17

335 T 335

seq\_name: gb\_est1:AU091413

seq\_documentation\_block: 399 bp mRNA EST 05-JUN-2000  
LOCUS AU091413 Rice cDNA from immature leaf including apical meristem  
DEFINITION AU091413 (under short day condition) Oryza sativa cDNA clone E60220, mRNA  
sequence.  
ACCESSION AU091413  
VERSION AU091413.1 GI:8251089  
KEYWORDS EST.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa

REFERENCE 1 (bases 1 to 399)  
Sasaki,T. and Yamamoto,K.  
Rice cDNA from immature leaf including apical meristem (2000)  
UNPUBLISHED (2000)  
CONTACT: Takuji Sasaki  
NATIONAL INSTITUTE OF AGRICULTURAL RESOURCES  
RICE GENOME RESEARCH PROGRAM, KANNONDAI 2-1-2, TSUKUBA, IBARAKI  
305-8602, JAPAN  
TEL: 81-298-38-7441  
FAX: 81-298-38-7468  
EMAIL: tsasaki@agr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT "RGF",  
E60220\_1A.

## FEATURES

source

Location/Qualifiers

1..399  
/organism="Oryza sativa"  
/strain="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="E60220"  
/clone\_lib="Rice cDNA from immature leaf including apical  
meristem (under short day condition)"  
/dev\_stage="immature leaf including apical meristem (under  
short day condition)"  
BASE COUNT 81 a 122 c 117 g 79 t  
ORIGIN

## alignment\_scores:

Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AU091413 ..

Align seg 1/1 to: AU091413 from: 1 to: 399

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
346 GACGAGTGTCTGGCGGCGCTCGGGTACAAAGTGGTGGTCCGACATGGC 395

17 a 17

396 C 396

seq\_name: gb\_est1:AL369482

seq\_documentation\_block: 458 bp mRNA EST 03-AUG-2000  
LOCUS AL369482 MtBA31D06F1 MtBA Medicago truncatula cDNA clone MtBA31D06 T3, mRNA  
DEFINITION AL369482  
sequence.  
ACCESSION AL369482  
VERSION AL369482.1 GI:9669235  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 458)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

Journet,E.P., Crespeau,H., van-Tuinen,D., Gouzy,J., Jaillon,O.,  
Niebel,A., Carreau,V., Chatagnier,O., Kahn,D., Gianinazzi-Pearson  
V. and Gamas,P.  
Medicago truncatula ESTs from nitrogen-starved roots  
UNPUBLISHED (2000)  
CONTACT: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

CONTACT: Pascal Gamas and Etienne-Pascal Journet, Laboratoire de  
Biologie Moleculaire des Relations Plantes-Microorganismes,  
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :  
Mt-est@toulouse.inra.fr Website :  
http://sequence.toulouse.inra.fr/Mtruncatula.html).

Location/Qualifiers

1..458  
/organism="Medicago truncatula"  
/cultivar="Jemalong"  
/db\_xref="taxon:3880"  
/clone="MtBA31D06"  
/clone\_lib="MtBA"  
/tissue\_type="root tips"  
/dev\_stage="harvested after 3 days of N-starvation"  
/note="Vector: pBluescript PSK; Site\_1: EcoRI; Site\_2:  
XhoI; Plants were grown in an aeroponic chamber for 14  
days on nitrogen-rich medium followed by 3 days on N-free  
medium. RNA was extracted from root tips (1-3 cm). cDNA  
was prepared from polyA+ enriched RNA. The cDNA was  
directionally ligated into Uni-zapXR vector from  
Stratagene and packaged using GigaPack Gold packaging  
extracts. Plasmids containing cDNA inserts were  
mass-excised from phage stocks using Exassit helper phage  
and propagated in SOLR cells. Clone ordering and  
sequencing was performed by the Centre National de  
Sequencage (Genoscope, Evry, France)."

BASE COUNT 172 a 97 c 82 g 107 t  
ORIGIN

## FEATURES

source

Location/Qualifiers

1..458  
/organism="Medicago truncatula"  
/cultivar="Jemalong"  
/db\_xref="taxon:3880"  
/clone="MtBA31D06"  
/clone\_lib="MtBA"  
/tissue\_type="root tips"  
/dev\_stage="harvested after 3 days of N-starvation"  
/note="Vector: pBluescript PSK; Site\_1: EcoRI; Site\_2:  
XhoI; Plants were grown in an aeroponic chamber for 14  
days on nitrogen-rich medium followed by 3 days on N-free  
medium. RNA was extracted from root tips (1-3 cm). cDNA  
was prepared from polyA+ enriched RNA. The cDNA was  
directionally ligated into Uni-zapXR vector from  
Stratagene and packaged using GigaPack Gold packaging  
extracts. Plasmids containing cDNA inserts were  
mass-excised from phage stocks using Exassit helper phage  
and propagated in SOLR cells. Clone ordering and  
sequencing was performed by the Centre National de  
Sequencage (Genoscope, Evry, France)."

BASE COUNT 172 a 97 c 82 g 107 t  
ORIGIN

## alignment\_scores:

Quality: 78.00 Length: 17  
Ratio: 4.588 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AL369482 ..

Align seg 1/1 to: AL369482 from: 1 to: 458

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
273 GACGAGTGTGTAGCAGCATAGGTACAAAGGTTTCGCTTCGACATGGC 322

17 a 17  
|  
323 T 323

seq\_name: gb\_est1:AL371425

seq\_documentation\_block: 466 bp mRNA EST 03-AUG-2000  
LOCUS AL371425  
DEFINITION MTBA Medicago truncatula cDNA clone MTBA44B05 T3, mRNA  
sequence.

ACCESSION AL371425  
VERSION AL371425.1 GI:9671178

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM

Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 466)

AUTHORS  
Journet,E.P., Crespeau,H., van-Tuinen,D., Gouzy,J., Jaillon,O.,  
Niebel,A., Carreau,V., Chatagnier,O., Kahn,D., Gianinazzi-Pearson  
,V. and Gamas,P.

TITLE Medicago truncatula ESTs from nitrogen-starved roots

JOURNAL Unpublished (2000)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de

Biologie Moleculaire des Relations Plantes-Microorganismes,

CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :

Mt-est@toulouse.inra.fr Website :

http://sequence.toulouse.inra.fr/Mtruncatula.html).

FEATURES

source

1. .466

/organism="Medicago truncatula"

/cultivar="Jemalong"

/db\_xref="taxon:3880"

/clone="MtBA44B05"

/clone\_lib="MtBA"

/tissue\_type="root tips"

/dev\_stage="harvested after 3 days of N-starvation"

/note="vector: pluescript psk; Site\_1: EcoRI; Site\_2:

XhoI; Plants were grown in an aeroponic chamber for 14

days on nitrogen-rich medium followed by 3 days on N-free

medium. RNA was extracted from root tips (1-3 cm). cDNA

was prepared from polyA+ enriched RNA. The cDNA was

directionally ligated into Uni-zapXR vector from

Stratagene and packaged using Gigapack Gold packaging

extracts. Plasmids containing cDNA inserts were

mass-excised from phage stocks using Exassit helper phage

and propagated in S04K cells. Clone ordering and

sequencing was performed by the Centre National de

Sequencage (Genoscope, Evry, France)."

BASE COUNT 183 a 93 c 82 g 108 t

ORIGIN

alignment\_scores:

Quality: 78.00 Length: 17

Ratio: 4.588 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x AL371425 ..

Align seg 1/1 to: AL371425 from: 1 to: 466

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17

|||||

283 GACGAGTGTGTACGACGATAGTACAGGTTCGCTCTCCGACATGCG 332

17 a 17  
|  
333 T 333

seq\_name: gb\_est1:AU222715

seq\_documentation\_block: 467 bp mRNA EST 30-JUL-2001  
LOCUS AU222715  
DEFINITION AU222715 Rice shoot Oryza sativa cDNA clone S0583, mRNA sequence.  
ACCESSION AU222715  
VERSION AU222715.1 GI:15008327

KEYWORDS EST.

SOURCE Oryza sativa.

ORGANISM

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoidae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 467)

AUTHORS Sasaki,T. and Yamamoto,K.

Rice cDNA from etiolated shoot (2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/

PROJECT "RGP"

S0583\_97A.

FEATURES Location/Qualifiers

source

1. .467

/organism="Oryza sativa"

/strain="Nipponbare, sub\_species Japonica"

/db\_xref="taxon:4530"

/clone="S0583"

/clone\_lib="Rice shoot"

/note="Etiolated shoot (8 days old)"

BASE COUNT 90 a 136 c 176 g 65 t

ORIGIN

alignment\_scores:

Quality: 78.00 Length: 17

Ratio: 4.588 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x AU222715 ..

Align seg 1/1 to: AU222715 from: 1 to: 467

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17

|||||

198 GACGAGTGTGTACGCGCTCGGTACAAGTTCGCTCGCATGCGC 247

17 a 17  
|  
248 C 248

seq\_name: gb\_est1:AV422153

seq\_documentation\_block:

LOCUS AV422153 484 bp mRNA EST 23-MAY-2000

DEFINITION AV422153 Lotus japonicus young plants (two-week old) Lotus

japonicus cDNA clone MWM005b04\_f 5', mRNA sequence.

ACCESSION AV422153

VERSION AV422153.1 GI:7776718

KEYWORDS EST.

SOURCE Lotus japonicus.

ORGANISM

Lotus japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

REFERENCE 1 (bases 1 to 484)  
 AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
 TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus  
 JOURNAL DNA Res. 7 (2), 127-130 (2000)  
 MEDLINE 20277479  
 COMMENT Contact: Yasukazu Nakamura  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: ynakamu@kazusa.or.jp. URL:http://www.kazusa.or.jp/en/plant/.

FEATURES  
 source  
 1..484  
 /organism="Lotus japonicus"  
 /db\_xref="taxon:34305"  
 /clone="MWM005b04\_r"  
 /clone\_lib="Lotus japonicus young plants (two-week old)"  
 /dev\_stage="young plants (two-week old)"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI; isolate=WiyakoJima MG-20".

BASE COUNT 130 a 144 c 103 g 107 t  
 ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:  
 US-09-485-529-104 x AV422153 ..

Align seg 1/1 to: AV422153 from: 1 to: 484

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 ||||||||||||||||||||||||||||||||||||||||||||  
 81 GATGAGGTTCTGCGGGTTTAGGTTACAAAGTTCCTCGCATGGC 130  
 17 a 17  
 131 T 131

seq\_name: gb\_estl:AW697326

seq\_documentation\_block:  
 LOCUS AW697326 537 bp mRNA EST 20-DEC-2000  
 DEFINITION NF117E10ST1F1082 Developing stem Medicago truncatula cDNA clone  
 ACCESSION AW697326  
 VERSION AW697326.2 GI:11933520  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 537)  
 AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
 ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon  
 ,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 JOURNAL Medicago truncatula stem library  
 COMMENT Unpublished (2000)  
 On Apr 14, 2000 this sequence version replaced gi:7572088.  
 Contact: Dixon RA  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7302  
 Fax: 580 221 7380

Email: radixon@noble.org  
 Insert Length: 686 Std Error: 0.00  
 Plate: 117 row: E column: 10  
 Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES  
 source  
 1..537  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /clone="NF117E10ST"  
 /clone\_lib="Developing stem"  
 /tissue\_type="stem"  
 /dev\_stage="pooled developmental"  
 /note="Vector: Lambda zap; Contains a mixture of  
 internodal stem segments".

BASE COUNT 199 a 120 c 87 g 131 t  
 ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:  
 US-09-485-529-104 x AW697326 ..

Align seg 1/1 to: AW697326 from: 1 to: 537

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 ||||||||||||||||||||||||||||||||||||||||||||  
 337 GACGAGTGTTCAGCATTTAGGTTACAAAGTTCGCTCCGACATGGC 386  
 17 a 17  
 387 T 387

seq\_name: gb\_estl:AW695914

seq\_documentation\_block:  
 LOCUS AW695914 559 bp mRNA EST 21-DEC-2000  
 DEFINITION NF099H09ST1F1079 Developing stem Medicago truncatula cDNA clone  
 ACCESSION AW695914  
 VERSION AW695914.2 GI:11957271  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 559)  
 AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
 ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon  
 ,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 JOURNAL Medicago truncatula stem library  
 COMMENT Unpublished (2000)  
 On Apr 14, 2000 this sequence version replaced gi:7570676.  
 Contact: Dixon RA  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7302  
 Fax: 580 221 7380

Email: radixon@noble.org  
 Insert Length: 654 Std Error: 0.00  
 Plate: 099 row: H column: 09  
 Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES  
 source  
 1..559  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"

/clone="NF099H05ST"  
 /clone\_lib="Developing stem"  
 /tissue\_type="stem"  
 /dev\_stage="Pooled developmental"  
 /note="Vector: Lambda zap; Contains a mixture of  
 intermodal stem segments"  
 BASE COUNT 197 a 121 c 83 g 158 t  
 ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118  
 alignment\_block:  
 US-09-485-529-104 x AW695914 ..  
 Align seg 1/1 to: AW695914 from: 1 to: 559

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 414 GACGAGTTGTTAGCAGCATAGTTACAAGGTTGCTCTTCGACATGGC 463

17 a 17  
 |  
 464 T 464

seq\_name: gb\_est2:BI308816

seq\_documentation\_block:  
 LOCUS BI308816 614 bp mRNA EST 20-JUL-2001  
 DEFINITION EST530226 GP0D Medicago truncatula cDNA clone pGP0D-8H2 5' end,  
 mRNA sequence.  
 ACCESSION BI308816  
 VERSION BI308816.1 GI:14983143  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 614)  
 Grusak, M.A., Sanac, D.A., Town, C.D., Van Aken, S., Utterback, T., Cho  
 , J., and Fraser, C.M.  
 ESTs from developing reproductive tissues of Medicago truncatula  
 Unpublished (2001)  
 Contact: Michael A. Grusak  
 USDA/ARS Children's Nutrition Research Center  
 Baylor College of Medicine  
 1100 Bates Street, Houston, TX 77030-2600, USA  
 Tel: 713-798-7044  
 Fax: 713-798-7078  
 Email: mgrusak@bcm.tmc.edu  
 B395671e  
 TIGR sequence name: WFOAX37TK  
 More information is available at: www.medicago.org  
 Seq primer: SKmod (CTA GAA CTA gtg gat CC).  
 Location/Qualifiers

FEATURES  
 source  
 1..614  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /cultivar="Al7"  
 /clone\_lib="Al7"  
 /clone="pgPOD-8H2"  
 /clone\_lib="GP0D"  
 /tissue\_type="immature pod walls"  
 /dev\_stage="immature pods, ranging in age from 15 to 30  
 days after pollination"  
 /note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Immature pods, ranging in age from 15 to 30 days  
 after pollination, were collected from greenhouse-grown  
 plants. At harvest, seeds were removed from pods and

isolated pod walls were collected and immediately frozen  
 in liquid nitrogen. Pod walls were pooled for mRNA  
 extraction. cDNA was prepared from polyA+ enriched RNA.  
 The cDNA was directionally ligated into the Unizap XR  
 vector from Stratagene and packaged using Gigapack III  
 Gold packaging extracts. Plasmids containing cDNA inserts  
 were excised from the recombinant lambda-zap phage using  
 Ex-assist helper phage and propagated in XL0LR cells."  
 BASE COUNT 220 a 130 c 114 g 150 t  
 ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

alignment\_block:

US-09-485-529-104 x BI308816 ..

Align seg 1/1 to: BI308816 from: 1 to: 614

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 236 GACGAGTTGTTAGCAGCATAGTTACAAGGTTGCTCTTCGACATGGC 285

17 a 17  
 |  
 286 T 286

seq\_name: gb\_est1:AW694061

seq\_documentation\_block:  
 LOCUS AW694061 655 bp mRNA EST 15-JUN-2000  
 DEFINITION NF072A05ST1F1036 Developing stem Medicago truncatula cDNA clone  
 NF072A05ST 5', mRNA sequence.  
 ACCESSION AW694061  
 VERSION AW694061.1 GI:7568798  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 655)  
 He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell  
 , C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon  
 , R.A.  
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 Medicago truncatula stem library  
 Unpublished (2000)  
 Contact: Dixon RA  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7302  
 Fax: 580 221 7380  
 Email: radixon@noble.org  
 Insert Length: 655 Std Error: 0.00  
 Plate: 072 row: A column: 05  
 Seq primer: TCACACAGGAACACAGCTATGAC.  
 Location/Qualifiers

FEATURES  
 source  
 1..655  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /clone\_lib="NF072A05ST"  
 /clone="NF072A05ST"  
 /tissue\_type="stem"  
 /dev\_stage="pooled developmental"  
 /note="vector: Lambda zap; Contains a mixture of  
 intermodal stem segments"  
 BASE COUNT 242 a 144 c 109 g 160 t

## ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AW694061 ..

Align seg 1/1 to: AW694061 from: 1 to: 655

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 298 GACGAGTTGTTAGCAGCATTAGGTTACAAGGTTTCGCTCTCCGACATGGC 347

17 a 17

348 T 348

seq\_name: gb\_est1:AW694064

seq\_documentation\_block: 656 bp mRNA EST 15-JUN-2000  
 LOCUS AW694064  
 DEFINITION NF072A07ST1052 Developing stem Medicago truncatula cDNA clone

ACCESSION AW694064

VERSION NF072A07ST 5', mRNA sequence.

## KEYWORDS

EST.

## SOURCE

barrel medic.

## ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 656)

He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell

,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon

,R.A.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula stem library

Unpublished (2000)

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 656 Std Error: 0.00

Plate: 072 row: A column: 07

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1..656

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF072A07St"

/clone\_lib="Developing stem"

/tissue\_type="stem"

/dev\_stage="Pooled developmental"

/note="Vector: Lambda Zap; Contains a mixture of

internodal stem segments"

BASE COUNT 238 a 145 c 107 g 166 t

ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AW694064 ..

Align seg 1/1 to: AW694064 from: 1 to: 656

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 305 GACGAGTTGTTAGCAGCATTAGGTTACAAGGTTTCGCTCTCCGACATGGC 354

17 a 17

355 T 355

seq\_name: gb\_est1:AW696350

seq\_documentation\_block:

LOCUS AW696350 667 bp mRNA EST 15-JUN-2000

DEFINITION NF107A12ST1088 Developing stem Medicago truncatula cDNA clone

NF107A12ST 5', mRNA sequence.

ACCESSION AW696350

VERSION AW696350.1 GI:7571200

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 667)

He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell

,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon

,R.A.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula stem library

Unpublished (2000)

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 667 Std Error: 0.00

Plate: 107 row: A column: 12

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1..667

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF107A12St"

/clone\_lib="Developing stem"

/tissue\_type="stem"

/dev\_stage="Pooled developmental"

/note="Vector: Lambda Zap; Contains a mixture of

internodal stem segments"

BASE COUNT 234 a 147 c 126 g 160 t

ORIGIN

alignment\_scores:  
 Quality: 78.00 Length: 17  
 Ratio: 4.588 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 94.118

## alignment\_block:

US-09-485-529-104 x AW696350 ..

Align seg 1/1 to: AW696350 from: 1 to: 667

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 254 GACGAGTTGTTAGCAGCATTAGGTTACAAGGTTTCGCTCTCCGACATGGC 303

17 a 17



304 T 304

seq\_name: gb\_est1:AW584593

```

seq_documentation_block:
LOCUS      AW584593      683 bp      mRNA      EST      07-SEP-2000
DEFINITION N210669e MHAM Medicago truncatula/Glomus versiforme mixed EST
library cDNA clone MHAM-2P17, mRNA sequence.
ACCESSION  AW584593
VERSION    AW584593.1 GI:7261647
KEYWORDS   EST.
SOURCE     Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM   Medicago truncatula/Glomus versiforme mixed EST library
Eukaryota; mixed EST libraries.
REFERENCE  1 (bases 1 to 683)
AUTHORS    Harrison,M.J., Liu,J., Peng,H., Gonzales,M., Ellis,L., Town,C.D.,
Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.
TITLE      ESTs from roots of Medicago truncatula after colonization with
Glomus versiforme
JOURNAL    Unpublished (2000)
COMMENT    Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Other name: MHAM-2c-H09; Date: 3/14/00; Updated to the Database of
Expressed Sequence Tags (dbEST) on 04/27/00; More information is
available at 'http://chryslie.tamu.edu/medicago'.
Seq primer: T3.

```

```

FEATURES             Location/Qualifiers
     source           1..683
                     /organism="Medicago truncatula/Glomus versiforme mixed EST
                     library"
                     /cultivar="Medicago truncatula genotype A17"
                     /db_xref="taxon:119092"
                     /clone="MHAM-2P17"
                     /clone_lib="MHAM"
                     /tissue_type="roots colonized with Glomus versiforme"
                     /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The library was
                     made from a mixture of RNA from each of these stages."
                     /lab_host="E. coli strain XL0LR"
                     /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The cDNA was
                     directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XL0LR cells."
BASE COUNT      244 a 152 c 114 g 173 t
ORIGIN

```

```

alignment_scores:
  Quality: 78.00      Length: 17
  Ratio: 4.588       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 94.118

```

```

alignment_block:
US-09-485-529-104 x AW584593

```

Align seg 1/1 to: AW584593 from: 1 to: 683

```

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
317 GACGAGTGTTCACGACATAGTTACAGGTCGCTCTCCGACATGCC 366

```

17 a 17

367 T 367

seq\_name: gb\_est1:AW584661

```

seq_documentation_block:
LOCUS      AW584661      701 bp      mRNA      EST      07-SEP-2000
DEFINITION N210765e MHAM Medicago truncatula/Glomus versiforme mixed EST
library cDNA clone MHAM-2P18, mRNA sequence.
ACCESSION  AW584661
VERSION    AW584661.1 GI:7261715
KEYWORDS   EST.
SOURCE     Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM   Medicago truncatula/Glomus versiforme mixed EST library
Eukaryota; mixed EST libraries.
REFERENCE  1 (bases 1 to 701)
AUTHORS    Harrison,M.J., Liu,J., Peng,H., Gonzales,M., Ellis,L., Town,C.D.,
Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.
TITLE      ESTs from roots of Medicago truncatula after colonization with
Glomus versiforme
JOURNAL    Unpublished (2000)
COMMENT    Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Other name: MHAM-2d-H09; Date: 3/14/00; Updated to the Database of
Expressed Sequence Tags (dbEST) on 04/27/00; More information is
available at 'http://chryslie.tamu.edu/medicago'.
Seq primer: T3.

```

```

FEATURES             Location/Qualifiers
     source           1..701
                     /organism="Medicago truncatula/Glomus versiforme mixed EST
                     library"
                     /cultivar="Medicago truncatula genotype A17"
                     /db_xref="taxon:119092"
                     /clone="MHAM-2P18"
                     /clone_lib="MHAM"
                     /tissue_type="roots colonized with Glomus versiforme"
                     /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The library was
                     made from a mixture of RNA from each of these stages."
                     /lab_host="E. coli strain XL0LR"
                     /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The cDNA was
                     directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XL0LR cells."
BASE COUNT      250 a 158 c 118 g 175 t
ORIGIN

```

```

alignment_scores:
  Quality: 78.00      Length: 17
  Ratio: 4.588       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 94.118

```

```

alignment_block:
US-09-485-529-104 x AW584661

```

Align seg 1/1 to: AW584661 from: 1 to: 701

```

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
317 GACGAGTGTTCACGACATAGTTACAGGTCGCTCTCCGACATGCC 366

```

17 a 17

```

367 T 367
seq_name: gb_estl:AA660952

seq_documentation_block:
LOCUS      AA660952      719 bp      mRNA      08-MAR-2000
DEFINITION 00849 MERHE Medicago truncatula cDNA 5', mRNA sequence.
ACCESSION  AA660952
VERSION    AA660952.1  GI:2604996
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 719)
AUTHORS    Covitz,P.A., Smith,L.S. and Long,S.R.
TITLE      Expressed sequence tags from a root-hair-enriched medicago
            truncatula cDNA library
JOURNAL    Plant Physiol. 117 (4), 1325-1332 (1998)
COMMENT    Contact: Long SR
            Department of Biological Sciences and Howard Hughes Medical
            Institute
            Stanford University
            Gilbert Biology, Stanford, CA 94305-5020, USA
            Tel: 650 723 3232
            Fax: 650 725 8309
            Email: fa.srl@leforsthe.stanford.edu
            Seq primer: T3.
            Location/Qualifiers
                1..719
                /organism="Medicago truncatula"
                /cultivar="Jemalong"
                /db_xref="taxon:3880"
                /clone_lib="MTRHE"
                /tissue_type="Root hairs & tips"
                /dev_stage="2-3 day old seedlings"
                /note="Organ: Root; Vector: pBK-CMV; Site_1: EcoRI;
                Site_2: XhoI; cDNA was synthesized from a pooled mRNA prep
                from elongating root hairs (30% w/w) and 2-3cm root tips
                (70% w/w). XhoI-oligo-dT linker-primer and EcoRI
                adaptors were used. cDNAs was cloned unidirectionally
                into lambda ZAP Express (Stratagene), amplified, and
                mass-exclised into pBK-CMV vector plasmids. More
                information is available at http://bio-srle8.stanford.edu."
BASE COUNT  255 a 163 c 122 g 172 t
ORIGIN
alignment_scores:
    Quality: 78.00      Length: 17
    Ratio: 4.588        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 94.118
alignment_block:
US-09-485-529-104 x AA660952
Align seg 1/1 to: AA660952 from: 1 to: 719
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
298 GACGAGTTGTTAGCAGCATAGGTTACAGGTTCCGCTCTCCGACATGCG 347
17 a 17
348 T 348
seq_name: gb_est2:BG647068
seq_documentation_block:
LOCUS      BG647068      748 bp      mRNA      EST      24-APR-2001
DEFINITION 00849 MERHE Medicago truncatula cDNA 5', mRNA sequence.
ACCESSION  BG647068
VERSION    BG647068.1  GI:13782180
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 748)
AUTHORS    Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
            Uterback,T., Cho,J. and Fraser,C.M.
TITLE      ESTs from roots of Medicago truncatula treated with
            oligogalacturonides of DP 6-20
JOURNAL    Unpublished (2001)
COMMENT    Contact: Michael G. Hahn
            Complex Carbohydrate Research Center
            University of Georgia
            220 Riverbend Road, Athens, GA 30602-4712, USA
            Tel: 706-542-4457
            Fax: 706-542-4412
            Email: hahn@ccrc.uga.edu
            G390574e TIGR sequence name: MTMBU48TK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
            Location/Qualifiers
                1..748
                /organism="Medicago truncatula"
                /cultivar="A17"
                /db_xref="taxon:3880"
                /clone_lib="phOGA-15H23"
                /tissue_type="HOGA"
                /dev_stage="24 hours after treatment in the dark at 26 C
                with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
                presence of 100 ug/ml Gentamicin"
                /lab_host="XLOLR"
                /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                was directionally ligated into the Unizap XR vector from
                Stratagene and packaged using Gigapack III Gold packaging
                extracts. Plasmids containing cDNA inserts were excised
                from the recombinant lambda-zap phage using Ex-assist
                helper phage and propagated in SOLR cells."
BASE COUNT  261 a 170 c 131 g 186 t
ORIGIN
alignment_scores:
    Quality: 78.00      Length: 17
    Ratio: 4.588        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 94.118
alignment_block:
US-09-485-529-104 x BG647068
Align seg 1/1 to: BG647068 from: 1 to: 748
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
327 GACGAGTTGTTAGCAGCATAGGTTACAGGTTCCGCTCTCCGACATGCG 376
17 a 17
377 T 377
seq_name: gb_est2:BG440209
seq_documentation_block:
LOCUS      BG440209      684 bp      mRNA      EST      15-MAR-2001
DEFINITION GA_Ea0006G06f Gossypium arboreum 7-10 dp fiber library Gossypium

```

arborescence cDNA clone GA\_Ea0006G06f, mRNA sequence.  
 BG440209  
 VERSION BG440209.1 GI:13349859  
 KEYWORDS EST.  
 SOURCE Gossypium arboreum.  
 ORGANISM Gossypium arboreum.

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

1 (bases 1 to 684)  
 Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.  
 TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: TAATACGACTCACTATAGG  
 High quality sequence stop: 682.

FEATURES  
 source  
 1..684  
 /organism="Gossypium arboreum"  
 /strain="AKA"  
 /cultivar="8400"  
 /db\_xref="taxon:29729"  
 /clone="GA\_Ea0006G06f"  
 /clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
 /tissue\_type="Fibers isolated from bolls harvested 7-10 dpa"  
 /lab\_host="E. coli"  
 /note="Vector: PBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 205 a 145 c 138 g 196 t  
 ORIGIN

alignment\_scores  
 Quality: 74.00 Length: 17  
 Ratio: 4.353 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 88.235

alignment\_block  
 US-09-485-529-104 x BG440209 ..

Align seg 1/1 to: BG440209 from: 1 to: 684

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 229 GACGAGTTTACGCTGTTTGGTTACAAAGTTCGGTCATCAGATATGCC 278

17 a 17  
 |  
 279 G 279

seq\_name: gb\_est1:AW584431

seq\_documentation\_block:  
 LOCUS AW584431 712 bp mRNA EST 07-SEP-2000  
 DEFINITION N210477E MHAM Medicago truncatula/Glomus versiforme mixed EST  
 library cDNA clone MHAM-2017, mRNA sequence.  
 ACCESSION AW584431  
 VERSION AW584431.1 GI:7261485  
 KEYWORDS EST.  
 SOURCE Medicago truncatula/Glomus versiforme mixed EST library.  
 ORGANISM Medicago truncatula/Glomus versiforme mixed EST library.  
 Eukaryota; mixed EST libraries.  
 1 (bases 1 to 712)  
 HARRISON, M.J., LIU, J., PENG, H., GONZALES, M., ELLIS, L., TOWN, C.D.,  
 BOWMAN, C.L., CRAVEN, M.B., HANSEN, T.S., HOLT, I.E. and FRASER, C.M.

ESTs from roots of Medicago truncatula after colonization with  
 Glomus versiforme  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Harrison M.J.  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: mjharrison@noble.org  
 Other name: MHAM-2a-H09; Date: 3/14/00; Updated to the Database of  
 Expressed Sequence Tags (dbEST) on 04/27/00; More information is  
 available at 'http://chryslie.tamu.edu/medicago'.  
 Seq primer: r3.

FEATURES  
 source  
 1..712  
 /organism="Medicago truncatula/Glomus versiforme mixed EST  
 library"  
 /cultivar="Medicago truncatula genotype A17"  
 /db\_xref="taxon:119092"  
 /clone="MHAM-2017"  
 /clone\_lib="MHAM"  
 /tissue\_type="roots colonized with Glomus versiforme"  
 /dev\_stage="Roots harvested at 10, 17, 22, 31 and 38 days  
 post-inoculation with Glomus versiforme. The library was  
 made from a mixture of RNA from each of these stages."  
 /lab\_host="E. coli strain XL0LR"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA from  
 roots harvested at 10, 17, 22, 31 and 38 days  
 post-inoculation with Glomus versiforme. The cDNA was  
 directionally ligated into the Unizap XR vector from  
 Stratagene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-zap phage using Ex-assist  
 helper phage and propagated in XL0LR cells."

BASE COUNT 249 a 160 c 128 g 175 t  
 ORIGIN

alignment\_scores  
 Quality: 74.00 Length: 17  
 Ratio: 4.353 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 88.235

alignment\_block  
 US-09-485-529-104 x AW584431 ..

Align seg 1/1 to: AW584431 from: 1 to: 712

1 AspGluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 317 GACGAGTTTACAGCATAGTATTAGTTACAGGTTCCGCTCTCCGACATGCC 366

17 a 17  
 |  
 367 T 367

seq\_name: gb\_est1:AW688657

seq\_documentation\_block:  
 LOCUS AW688657 651 bp mRNA EST 15-JUN-2000  
 DEFINITION NF010B04ST1F1000 Developing stem Medicago truncatula cDNA clone  
 NF010B04ST 5', mRNA sequence.  
 ACCESSION AW688657  
 VERSION AW688657.1 GI:7563393  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

```

REFERENCE 1 (bases 1 to 651)
AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
          C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
          R.A.
TITLE      Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL    Medicago truncatula stem library
COMMENT     Unpublished (2000)
          Contact: Dixon RA
          Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7302
          Fax: 580 221 7380
          Email: radixon@noble.org
          Insert Length: 651 Std Error: 0.00
          Plate: 010 row: B column: 04
          Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES   Location/Qualifiers
            1..651
             /organism="Medicago truncatula"
             /db_xref="taxon:3880"
             /clone_lib="NF010B04ST"
             /clone_lib="developing stem"
             /tissue_type="stem"
             /dev_stage="Pooled developmental"
             /note="Vector: Lambda Zap; Contains a mixture of
             internodal stem segments"
BASE COUNT 193 a 139 c 143 g 175 t 1 others
ORIGIN
alignment_scores:
  Quality: 72.00 Length: 16
  Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 93.750
alignment_block:
US-09-485-529-104 x AW688657 ..
Align seg 1/1 to: AW688657 from: 1 to: 651
2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
50 GAGTGTGTAGCAGCATAGGTACAAAGTTCGCTCCGACATGGCT 97
seq_name: gb_est1:AI495884
seq_documentation_block:
LOCUS      AI495884 311 bp mRNA EST 01-DEC-1999
DEFINITION sb17d07.y1 Gm-cl004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-cl004-9062 5' similar to TR:O23643 O23643 RGA2 PROTEIN. [1] ;,
            mRNA sequence.
ACCESSION  AI495884
VERSION    AI495884.1 GI:4396887
KEYWORDS   EST.
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE  1 (bases 1 to 311)
AUTHORS   Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
          A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
          Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          R., Waterston,R. and Willson,R.
            Public Soybean EST Project
            Unpublished (1999)
            Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine

```

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Gibco
High quality sequence stop: 307.
FEATURES   Location/Qualifiers
            1..311
             /organism="Glycine max"
             /db_xref="taxon:3847"
             /clone="GENOME SYSTEMS CLONE ID: Gm-cl004-9062"
             /clone_lib="Gm-cl004"
             /tissue_type="root"
             /lab_host="XL10-Gold"
             /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
             XhoI; Root cDNA. The mRNA was isolated from entire roots
             of 8 day old 'Williams' seedlings which were propagated on
             paper towels with distilled water. Stratagene's cDNA
             Synthesis Kit (catalog #200401) was used to synthesize the
             cDNA. First- strand synthesis was performed with 5-methyl
             dCTP, hence the ligated cDNA is hemimethylated.
             Stratagene's first-strand synthesis primer was used
             [GAGAGAGAGAGAGAGAGACTAGTCGAG(T)-18]. After
             second-strand synthesis, the cDNA ends were 'polished'
             with clone Pfu DNA polymerase, ligated to EcoRI adapters,
             and phosphorylated. The XhoI site within the first-strand
             synthesis primer was restricted by digestion with XhoI;
             all XhoI sites in the cDNA would be protected by their
             hemimethylated status. The cDNA constructs were
             size-fractionated with a 500bp cutoff, using GibcoBRL Life
             Technologies' cDNA Size Fractionation column. The column
             eluent was then ligated into Stratagene's pBluescript II
             XR Predigested vector (pBluescript II SK(+)) that had been
             digested with EcoRI and XhoI, and phosphorylated). Both
             the white and blue colonies appear to contain recombinant
             plasmids with cDNA inserts. Blue colonies 9n-15) have been
             sequenced, and possess putative cDNA inserts. This library
             was constructed by Dr. Paul Kelm & Virginia H. Coryell,
             Department of Biology, Box5640, Northern Arizona
             University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
             Paul Kelm), 520-523-1372 (Virginia H. Coryell), Fax:
             520-523-7500, email: paul.keim@nau.edu,
             virginia.coryell@nau.edu"
BASE COUNT 87 a 74 c 88 g 62 t
ORIGIN
alignment_scores:
  Quality: 71.00 Length: 17
  Ratio: 4.176 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353
alignment_block:
US-09-485-529-104 x AI495884 ..
Align seg 1/1 to: AI495884 from: 1 to: 311
1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
115 GACGAGCTCTTCGAGTGGTGGGTACAAAGGTGAGGTGCATCGACATGCG 164
17 a 17
165 G 165
seq_name: gb_est1:AV409731
seq_documentation_block:

```

LOCUS AV409731 365 bp mRNA EST 23-MAY-2000  
 DEFINITION AV409731 Lotus japonicus young plants (two-week old) Lotus japonicus cDNA clone MWL061e11\_r 5', mRNA sequence.  
 ACCESSION AV409731  
 VERSION AV409731.1 GI:7722585  
 KEYWORDS EST.  
 SOURCE Lotus japonicus.  
 ORGANISM Lotus japonicus.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.  
 REFERENCE 1 (bases 1 to 365)  
 AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
 TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus  
 JOURNAL DNA Res. 7 (2), 127-130 (2000)  
 MEDLINE 20277479  
 COMMENT Contact: Yasukazu Nakamura  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: ynakam@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.  
 FEATURES  
 source  
 1. 365  
 /organism="Lotus japonicus"  
 /db\_xref="taxon:34305"  
 /clone="MWL061e11\_r"  
 /clone\_lib="Lotus japonicus young plants (two-week old)"  
 /stage="young plants (two-week old)"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI; Isolate=Miyakojima MG-20"  
 XhoI  
 106 a 95 c 102 g 62 t  
 BASE COUNT  
 ORIGIN  
 alignment\_scores  
 Quality: 71.00 Length: 17  
 Ratio: 4.176 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 82.353  
 alignment\_block  
 US-09-485-529-104 x AV409731 ..  
 Align seg 1/1 to: AV409731 from: 1 to: 365  
 1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
 |||||  
 118 GAGGAGCTTCGCGGTGGTGGTTACAAGGTGAGGTGATCATGACATGGC 167  
 17 a 17  
 168 G 168  
 seq\_name: gb\_est1:BE659954  
 seq\_documentation\_block  
 LOCUS BE659954 455 bp mRNA EST 06-SEP-2000  
 DEFINITION 746 GmaxSC Glycine max cDNA, mRNA sequence.  
 ACCESSION BE659954  
 VERSION BE659954.1 GI:9985948  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 REFERENCE 1 (bases 1 to 455)  
 AUTHORS Harris,N., Chapman,B.P. and Gijzen,M.  
 TITLE Gene expression in developing soybean seed coats  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Gijzen M

Agriculture and Agri-Food Canada  
 1391 Sandford Street, London, Ontario, Canada N5V 4T3  
 Tel: 519 457 1470  
 Fax: 519 457 3997  
 Email: gijzenm@agr.ca.  
 Location/Qualifiers  
 1. 455  
 /organism="Glycine max"  
 /cultivar="Harosoy 63"  
 /db\_xref="taxon:3847"  
 /clone\_lib="GmaxSC"  
 /tissue\_type="Seed coats"  
 /lab\_host="E. coli strain XL0LR"  
 /note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI; This cDNA library was constructed from polyA+ enriched mRNA from green seed coats in mid to late developmental stage, average fresh weight 250 mg per seed. Traces of pod and embryo tissue also present. Complementary DNA was synthesized from mRNA using an XhoI-poly(dT) linker-primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments and the products were digested with XhoI for directional cloning into lambda ZAP Express vector. This lambda library was amplified once using E. coli host strain XL1 Blue MRF'. Inserts were then subcloned by mass excision using ExAssist helper phage for conversion into phagemid vector pBK-CMV in E. coli host strain XL0LR."  
 104 a 115 c 113 g 122 t 1 others  
 BASE COUNT  
 ORIGIN  
 alignment\_scores  
 Quality: 71.00 Length: 16  
 Ratio: 4.438 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 93.750  
 alignment\_block  
 US-09-485-529-104 x BE659954 ..  
 Align seg 1/1 to: BE659954 from: 1 to: 455  
 2 GluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17  
 ::::|  
 332 AAGTTGCTGGCGGTGGGTACAAGGTTCGTGCTCCGACATGCC 379  
 seq\_name: gb\_est2:BF424878  
 seq\_documentation\_block  
 LOCUS BF424878 465 bp mRNA EST 28-NOV-2000  
 DEFINITION su52b08.y1 Gm-cl069 Glycine max cDNA clone GENOME SYSTEMS. CLONE ID: Gm-cl069-471 5' similar to TR:023642 O23642 RCAL PROTEIN. [1] ; mRNA sequence.  
 ACCESSION BF424878  
 VERSION BF424878.1 GI:11412867  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 REFERENCE 1 (bases 1 to 465)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swallier,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine,

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com

High quality sequence stop: 412.

#### FEATURES

source  
1. .465  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl069-471"  
/clone\_lib="Gm-cl069"  
/tissue\_type="degenerating cotyledons, 9-10 day old etiolated seedling"  
/lab\_host="DH10B"

/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from degenerating cotyledons of 9-10 day old etiolated seedlings for the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 120 a 116 c 137 g 92 t  
ORIGIN

#### alignment\_scores:

Quality: 71.00 Length: 17  
Ratio: 4.176 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

#### alignment\_block:

US-09-485-529-104 x BF424878 ..

Align seg 1/1 to: BF424878 from: 1 to: 465

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
114 GACGAGCTTCTCGGTTGAGTTACAGGTTAGTTCATCGGACATGCG 163

17 a 17

164 G 164

#### seq\_name: gb\_est1:AW720697

seq\_documentation\_block: 478 bp mRNA EST 19-APR-2000  
LOCUS AW720697  
DEFINITION LjNEST/h8rc Lotus japonicus nodule library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.

ACCESSION AW720697

VERSION AW720697.1 GI:7615248

KEYWORDS EST.

SOURCE Lotus japonicus.

ORGANISM Lotus japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

REFERENCE 1 (bases 1 to 478)

AUTHORS Colebatch,G., Freund,S., Trevaskis,B and Udvardi,M.

TITLE Lotus japonicus root nodule ESTs: tools for functional genomics

JOURNAL Unpublished (2000)

#### COMMENT

Contact: Udvardi MK  
Molecular Plant Nutrition  
Max Planck Institute of Molecular Plant Physiology  
Am Muehlenberg 1, 14476 Golm, Germany  
Fax: 49 331 567 8250  
Email: udvardi@mpimp-golm.mpg.de  
Seq primer: T7  
High quality sequence stop: 478.

#### FEATURES

source  
1. .478  
/organism="Lotus japonicus"  
/cultivar="Gifu (B-129)"  
/db\_xref="taxon:34305"  
/clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
/dev\_stage="5 and 7 week-old plants"  
/note="Organ: Nodule; Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

BASE COUNT 152 a 127 c 122 g 77 t  
ORIGIN

#### alignment\_scores:

Quality: 71.00 Length: 17  
Ratio: 4.176 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

#### alignment\_block:

US-09-485-529-104 x AW720697 ..

Align seg 1/1 to: AW720697 from: 1 to: 478

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
172 GACGAGCTTCTCGGTTGAGTTACAGGTTAGTTCATCGACATGCG 221

17 a 17

222 G 222

#### seq\_name: gb\_est2:BI419686

#### seq\_documentation\_block:

LOCUS BI419686 586 bp mRNA EST 15-AUG-2001  
DEFINITION LjNEST47e12r Lotus japonicus nodule library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.

ACCESSION BI419686

VERSION BI419686.1 GI:15190709

KEYWORDS EST.

SOURCE Lotus japonicus.

ORGANISM Lotus japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

REFERENCE 1 (bases 1 to 586)

AUTHORS Colebatch,G., Freund,S., Trevaskis,B and Udvardi,M.

TITLE Lotus japonicus root nodule ESTs: tools for functional genomics

JOURNAL Unpublished (2000)

COMMENT Contact: Udvardi MK

Molecular Plant Nutrition

Max Planck Institute of Molecular Plant Physiology

Am Muehlenberg 1, 14476 Golm, Germany

Fax: 49 331 567 8250

Email: udvardi@mpimp-golm.mpg.de

Seq primer: T7

High quality sequence stop: 586.

Location/Qualifiers

1. .586

/organism="Lotus japonicus"

```

/cultivar="Gifu (B-129)"
/db_xref="taxon:34305"
/clone_lib="Lotus japonicus nodule library 5 and 7
week-old"
/dev_stage="5 and 7 week-old plants"
/notes="Organ: Nodule; Vector: pSPORT1; Site_1: Salt;
Site_2: NotI; The library was prepared using mRNA
extracted from nodules of 5 and 7 week-old Lotus plants.
Nodules were induced by, and contained Mesorhizobium
strain R7A."

```

BASE COUNT 149 a 173 c 161 g 103 t

alignment\_scores:  
 Quality: 71.00 Length: 17  
 Ratio: 4.176 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 82.353

alignment\_block:

US-09-485-529-104 x BI419686 ..

Align seg 1/1 to: BI419686 from: 1 to: 586

```

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
119 GACGAGCTCTCGCGGTGGTGGGTACAGGTGAGGTGATCAGCATGGC 168

```

17 a 17

1

169 G 169

seq\_name: gb\_est1:AW267879

```

seq_documentation_block:
LOCUS AW267879 587 bp mRNA EST 07-SEP-2000
DEFINITION EST306157 DSIR Medicago truncatula cDNA clone pDSIR-8K8, mRNA
sequence.
ACCESSION AW267879.1 GI:6654835
VERSION AW267879
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

```

```

REFERENCE
1 (bases 1 to 587)
Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.
ESTs from roots of Medicago truncatula after inoculation with
Phytophthora medicaginis
Unpublished (1999)
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
Minnesota EST name:M250388e
TIGR sequence name:MTBAF64PK
More information, including clone ordering, is available at.
http://chrystle.tamu.edu/medicago
Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
Location/Qualifiers
1. 587

```

```

FEATURES
source
/cultivar="Medicago truncatula"
/cultivar="genotype Al7"
/db_xref="taxon:3880"
/clone="pDSIR-8K8"
/clone_lib="DSIR"

```

```

/tissue_type="roots infected with Phytophthora
medicaginis"
/dev_stage="roots harvested at 10 days post inoculation
with Phytophthora medicaginis"
/lab_host="E. coli strain XL0LR"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
roots harvested at 10 days post inoculation with
Phytophthora medicaginis. The cDNA was directionally
ligated into the Uni-ZAP XR vector from Stratagene and
packaged using Gigapack III Gold packaging extracts. The
Plasmids containing cDNA inserts were excised from the
recombinant lambda-ZAP phage using Ex-Assist helper phage
and propagated in XL0LR cells. Note: EST may be of fungal
origin."

```

BASE COUNT 168 a 129 c 127 g 163 t

ORIGIN

alignment\_scores:

Quality: 71.00 Length: 17  
 Ratio: 4.438 Gaps: 0  
 Percent Similarity: 94.118 Percent Identity: 88.235

alignment\_block:

US-09-485-529-104 x AW267879 ..

Align seg 1/1 to: AW267879 from: 1 to: 587

```

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17
|||||
3 GACGAGTTGTTAGCAGCATTAGGTTACAGGTTCTCTCTCCGACATGGC 52

```

17 a 17

1

53 T 53

seq\_name: gb\_est1:BE524828

```

seq_documentation_block:
LOCUS BE524828 374 bp mRNA EST 19-MAR-2001
DEFINITION M54G7STM Arabidopsis developing seed Arabidopsis thaliana cDNA
clone M54G7 5', mRNA sequence.
ACCESSION BE524828
VERSION BE524828.1 GI:9782806
KEYWORDS EST.
SOURCE thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

```

```

REFERENCE
1 (bases 1 to 374)
White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
Plant Physiol. 124 (4), 1582-1594 (2000)
MEDLINE
20567808
COMMENT
Contact: Benning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824
, USA
Tel: 517 355 1609
Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsids
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
6142920603 TEL: 6142929371.
Location/Qualifiers
1. 374
/organism="Arabidopsis thaliana"

```

```

/strain="Columbia"
/db_xref="taxon:3702"
/clone="M54G7"
/clone_lib="Arabidopsis developing seed"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/note="Organ: Developing seed; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoII"
BASE COUNT 97 a 88 c 87 g 102 t
ORIGIN

alignment_scores:
  Quality: 70.00 Length: 17
  Ratio: 4.118 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x BE524828 ..
Align seg 1/1 to: BE524828 from: 1 to: 374
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
100 GATGAGCTTCTGCTGTCTTGGCTACAAAGGTCGATCTTCTGAGATGCC 149

17 a 17
|
150 T 150

seq_name: gb_est2:T22782

seq documentation_block:
LOCUS T22782 457 bp mRNA EST 06-NOV-1997
DEFINITION 4790 Lambda-PRL2 Arabidopsis thaliana cDNA clone 107E8T7, mRNA
sequence.
ACCESSION T22782
VERSION T22782.1 GI:2597312
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 457) Plantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
AUTHORS Newman,F., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlroge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL 95148729
MEDLINE
COMMENT On Nov 6, 1997 this sequence version replaced gi:932620.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
  source
    Location/Qualifiers
      1..457
        /organism="Arabidopsis thaliana"
        /strain="var columbia"
        /db_xref="taxon:3702"
        /clone="107E8T7"
        /clone_lib="Lambda-PRL2"
        /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
        Lambda PRL2 is a cDNA library derived from equal
        quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 141 a 86 c 87 g 125 t 18 others
ORIGIN

alignment_scores:
  Quality: 70.00 Length: 17
  Ratio: 4.118 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 82.353

alignment_block:
US-09-485-529-104 x T22782 ..
Align seg 1/1 to: T22782 from: 1 to: 457
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
|||||
256 GATGAGCTTCTGCTGTCTTGGTATAAGGTTAGGTCATCCGAATGCC 305

17 a 17
|
306 T 306

seq_name: gb_est1:AV526467

seq documentation_block:
LOCUS AV526467 545 bp mRNA EST. 01-SEP-2000
DEFINITION AV526467 Arabidopsis thaliana aboveground organs two to six-week
old Arabidopsis thaliana cDNA clone APZ14e05R 5', mRNA sequence.
ACCESSION AV526467
VERSION AV526467.1 GI:8685995
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 545)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL 20363093
MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
  source
    Location/Qualifiers
      1..545
        /organism="Arabidopsis thaliana"
        /strain="Columbia"
        /db_xref="taxon:3702"
        /clone="APZ14e05R"
        /clone_lib="Arabidopsis thaliana aboveground organs two to
        six-week old"
        /tissue_type="aboveground organs"
        /dev_stage="two to six-week old"
        /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
        XhoI"
BASE COUNT 151 a 120 c 132 g 142 t
ORIGIN

alignment_scores:
  Quality: 70.00 Length: 17

```



Ratio: 4.118 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 82.353

## alignment\_block:

US-09-485-529-104 x AV526467 ..

Align seg 1/1 to: AV526467 from: 1 to: 545

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||  
188 GATGAGCTTCTTCTGCTCTTGGCTACAAAGGTCGATCTCTGAGATGCG 237

17 a 17

238 T 238

seq\_name: gb\_est2:BG442853

seq\_documentation\_block: 655 bp mRNA EST 15-MAR-2001  
LOCUS BG442853  
DEFINITION GA\_Ea0018120f Gossypium arboreum 7-10 dpa fiber library Gossypium  
arboreum cDNA clone GA\_Ea0018120f, mRNA sequence.  
ACCESSION BG442853  
VERSION BG442853.1 GI:13352505  
KEYWORDS EST.  
SOURCE Gossypium arboreum.  
ORGANISM Gossypium arboreum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 655)  
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry  
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

TITLE An integrated analysis of the genetics, development, and evolution  
of the cotton fiber  
JOURNAL Unpublished (2000)

COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

Seq primer: TAATACGACTCACTATAGG

High quality sequence stop: 647.

## FEATURES

source

1..655

/organism="Gossypium arboreum".

/strain="AKA"

/cultivar="8400"

/db\_xref="taxon:29729"

/clone="GA\_Ea0018120f"

/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"

/tissue\_type="Fibers isolated from bolls harvested 7-10

dpa"

/lab\_host="E. coli"

/note="vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT

165 a 188 c 132 g 170 t

## alignment\_scores:

Quality: 70.00 Length: 17

Ratio: 4.118 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 82.353

## alignment\_block:

US-09-485-529-104 x BG442853 ..

Align seg 1/1 to: BG442853 from: 1 to: 655

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||

195 GATGAGCTTTTGGCGGTTTTGGTTTACAAGTCAAAACTTCAGACATGCG 244

17 a 17

245 T 245

seq\_name: gb\_est2:BI267918

seq\_documentation\_block: 227 bp mRNA EST 18-JUL-2001  
LOCUS BI267918  
DEFINITION NF112E07IN1F1055 Insect herbivory Medicago truncatula cDNA clone  
NF112E07IN 5', mRNA sequence.

ACCESSION BI267918

VERSION BI267918.1 GI:14873367

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 227)

AUTHORS Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores  
,H.R., Inman,J.T., Weiler,J.W. and May,G.D.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL Medicago truncatula insect herbivory library

COMMENT Unpublished (2000)

Contact: Korth K

Dept. of Plant Pathology

University of Arkansas

217 Plant Science Building, Fayetteville, AR 72701, USA

Tel: 501 575 5191

Fax: 501 575 7601

Email: kthorth@comp.uark.edu

Insert Length: 227 Std Error: 0.00

Plate: 112 row: E column: 07

Seq primer: TCACACAGGAACACGCTATGAC.

## FEATURES

source

1..227

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF112E07IN"

/clone\_lib="insect herbivory"

/tissue\_type="local and systemic leaves"

/dev\_stage="mature"

/note="vector: Lambda Zap; Library was produced from fully  
expanded M. truncatula leaves of plants fed upon by  
Spodoptera exigua (beet armyworm) for 24 hours. Systemic  
(undamaged leaves from injured plants) and wounded leaves  
were harvested and pooled."

BASE COUNT 74 a 27 c 67 g 57 t 2 others

ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17

Ratio: 4.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x BI267918 ..

Align seg 1/1 to: BI267918 from: 1 to: 227

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
|||||

144 GATGAACTCTTAGCTGTAGTTGGTTACAAGTGAATCTTCAGACATGCG 193

17 a 17

194 T 194

seq\_name: gb\_est1:BE317531

seq\_documentation\_block: 288 bp mRNA EST 21-DEC-2000  
 LOCUS BE317531  
 DEFINITION NF051B06LFF1045 Developing leaf Medicago truncatula cDNA clone  
 ACCESSION BE317531  
 VERSION BE317531.2 GI:11961937  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 REFERENCE 1 (bases 1 to 288)  
 AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 JOURNAL Medicago truncatula leaf library  
 COMMENT Unpublished (2000)  
 On Jul 14, 2000 this sequence version replaced gi:9191308.

CONTACT: May GD  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391  
 Fax: 580 221 7380  
 Email: gdmay@noble.org  
 Medicago Genome Initiative accession: MGI:S:21598  
 Insert Length: 766 Std Error: 0.00  
 Plate: 051 row: B column: 06  
 Seq primer: TCACACAGGAACAGCTATGAC.

## FEATURES

source

1..288  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /clone="NF051B06LF"  
 /clone\_lib="Developing leaf"  
 /tissue\_type="leaf"  
 /dev\_stage="Pooled developmental"  
 /note="Vector: Lambda Zap; Contains a mixture of very  
 young, developing, mature and senescing leaves."

BASE COUNT 107 a 34 c 77 g 70 t  
 ORIGIN

alignment\_scores:  
 Quality: 68.00 Length: 17  
 Ratio: 4.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 76.471

alignment\_block:  
 US-09-485-529-104 x BE317531 ..

Align seg 1/1 to: BE317531 from: 1 to: 288

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 175 GATGAACTCTTAGCTAGTGGTTACAAAGTGAATCTTCAGACATGGC 224

17 a 17  
 225 T 225

seq\_name: gb\_est1:AL371416

seq\_documentation\_block:  
 LOCUS AL371416 398 bp mRNA EST 03-AUG-2000  
 DEFINITION MTBA44A11F1 MtBA Medicago truncatula cDNA clone MTBA44A11 T3, mRNA  
 sequence.  
 ACCESSION AL371416  
 VERSION AL371416.1 GI:9671169  
 KEYWORDS EST.

## SOURCE

ORGANISM

barrel medic.  
 Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 REFERENCE 1 (bases 1 to 398)  
 AUTHORS Journet, E.P., Crespeau, H., van Tuinen, D., Gouzy, J., Jaillon, O.,  
 Niebel, A., Carreau, V., Chatagnier, O., Kahn, D., Gianinazzi-Pearson  
 V. and Gamas, P.

## TITLE

JOURNAL

COMMENT

Medicago truncatula ESTs from nitrogen-starved roots

Unpublished (2000)  
 Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Contact: Pascal Gamas and Etienne-Pascal Journet, Laboratoire de

Biologie Moleculaire des Relations Plantes-Microorganismes,

CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email: [MT-est@toulouse.inra.fr](mailto:MT-est@toulouse.inra.fr) Website: <http://sequence.toulouse.inra.fr/Mtruncatula.html>).

Location/Qualifiers

1..398  
 /organism="Medicago truncatula"  
 /cultivar="Jemalong"  
 /db\_xref="taxon:3880"  
 /clone="MTBA44A11"  
 /clone\_lib="MtBA"  
 /tissue\_type="root tips"  
 /dev\_stage="harvested after 3 days of N-starvation"  
 /note="Vector: pBluescript PSK; Site 1: EcoRI; Site 2:  
 XhoI; Plants were grown in an aeroponic chamber for 14  
 days on nitrogen-rich medium followed by 3 days on N-free  
 medium. RNA was extracted from root tips (1-3 cm). cDNA  
 was prepared from polyA+ enriched RNA. The cDNA was  
 directionally ligated into Uni-zapXR vector from  
 Stratagene and packaged using Gigapack Gold packaging  
 extracts. Plasmids containing cDNA inserts were  
 mass-excised from phage stocks using ExAssit helper phage  
 and propagated in SOLR cells. Clone ordering and  
 sequencing was performed by the Centre National de  
 Sequencage (Genoscope, Evry, France)."

BASE COUNT 140 a 69 c 88 g 101 t  
 ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17  
 Ratio: 4.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x AL371416 ..

Align seg 1/1 to: AL371416 from: 1 to: 398

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
 |||||  
 178 GATGAACTCTTAGCTAGTGGTTACAAAGTGAATCTTCAGACATGGC 227

17 a 17  
 228 T 228

seq\_name: gb\_est2:BI267177

seq\_documentation\_block:  
 LOCUS BI267177 442 bp mRNA EST 18-JUL-2001  
 DEFINITION NF102E07IN1F1055 Insect herbivory Medicago truncatula cDNA clone  
 sequence.  
 ACCESSION BI267177  
 VERSION BI267177.1 GI:14871977  
 KEYWORDS EST.

SOURCE  
ORGANISM  
barrel medic.  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE  
1 (bases 1 to 442)  
AUTHORS  
Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores  
H.R., Inman, J.T., Weller, J.W. and May, G.D.  
TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
JOURNAL  
Medicago truncatula insect herbivory library  
COMMENT  
Unpublished (2000)  
Dept. of Plant Pathology  
University of Arkansas  
217 Plant Science Building, Fayetteville, AR 72701, USA  
Tel: 501 575 5191  
Fax: 501 575 7601  
Email: korth@comp.uark.edu  
Insert Length: 442 Std Error: 0.00  
Plate: 102 row: E column: 07  
Seq primer: TCACACAGGAACAGCTATGAC.  
FEATURES  
source  
Location/Qualifiers  
1..442  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF102E07IN"  
/clone\_lib="Insect herbivory"  
/tissue\_type="local and systemic leaves"  
/dev\_stage="mature"  
/note="Vector: Lambda Zap; Library was produced from fully  
expanded M. truncatula leaves of plants fed upon by  
Spodoptera exigua (beet armyworm) for 24 hours. Systemic  
(undamaged leaves from injured plants) and wounded leaves  
were harvested and pooled."  
BASE COUNT 154 a 81 c 93 g 112 t 2 others  
ORIGIN

alignment\_scores:  
Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

alignment\_block:  
US-09-485-529-104 x BI267177 ..  
17 a 17  
236 T 236

Align seg 1/1 to: BI267177 from: 1 to: 442  
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
186 GATGAACCTCTAGCTAGTGTGTTACAAAGTGAAATCTTCAGACATGCC 235  
17 a 17  
236 T 236

seq\_name: gb\_est2:BF634231

seq\_documentation\_block:  
LOCUS BF634231 454 bp mRNA EST 19-DEC-2000  
DEFINITION :NF084H12DTrF1103 Drought Medicago truncatula cDNA clone NF084H12DT  
5', mRNA sequence.  
ACCESSION :BF634231  
VERSION :BF634231.1 GI:11898389  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE  
1 (bases 1 to 454)

AUTHORS  
Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
JOURNAL  
Medicago truncatula drought library  
COMMENT  
Unpublished (2000)  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 454 Std Error: 0.00  
Plate: 084 row: H column: 12  
Seq primer: TCACACAGGAACAGCTATGAC.  
FEATURES  
source  
Location/Qualifiers  
1..454  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF084H12DT"  
/clone\_lib="Drought"  
/tissue\_type="Plantlets"  
/dev\_stage="pooled timepoints"  
/note="Vector: Lambda Zap; Contains a mixture of entire  
plantlets harvested in a series of days-post-watering  
timepoints."  
BASE COUNT 150 a 92 c 96 g 115 t 1 others  
ORIGIN

alignment\_scores:  
Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

alignment\_block:  
US-09-485-529-104 x BF634231 ..  
17 a 17  
208 T 208

Align seg 1/1 to: BF634231 from: 1 to: 454  
1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAl 17  
158 GATGAACCTCTAGCTAGTGTGTTACAAAGTGAAATCTTCAGACATGCC 207  
17 a 17  
208 T 208

seq\_name: gb\_est1:AW691269

seq\_documentation\_block:  
LOCUS AW691269 470 bp mRNA EST 21-DEC-2000  
DEFINITION :NF039G01STrF1000 Developing stem Medicago truncatula cDNA clone  
NF039G01ST 5', mRNA sequence.  
ACCESSION :AW691269  
VERSION :AW691269.2 GI:11957372  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE  
1 (bases 1 to 470)  
AUTHORS  
He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell  
C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon  
R.A.  
TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
JOURNAL  
Medicago truncatula stem library  
COMMENT  
Unpublished (2000)  
On Apr 14, 2000 this sequence version replaced gi:7566005.  
Contact: Dixon RA  
Plant Biology Division

The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7302  
Fax: 580 221 7380  
Email: radixon@noble.org

Insert Length: 678 Std Error: 0.00  
Plate: 039 row: G column: 01  
Seq primer: TCACACAGGAACACGCTATGAC.

## FEATURES

source  
1. 470  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone\_lib="NF039G01St"  
/clone\_lib="Developing stem"  
/tissue\_type="stem"  
/dev\_stage="Pooled developmental"  
/note="Vector: Lambda Zap; Contains a mixture of  
internodal stem segments"  
BASE COUNT 151 a 98 c 98 g 123 t  
ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x AW691269 ..  
Align seg 1/1 to: AW691269 from: 1 to: 470  
1 AspCluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
160 GATGAACCTCTAGCTAGTGGTTACAAAAGTGAATCTTCAGACATGGC 209

17 a 17

210 T 210

seq\_name: gb\_est1:BE202661

## seq\_documentation\_block:

LOCUS BE202661 482 bp mRNA EST 07-SEP-2000  
DEFINITION EST402683 KV1 Medicago truncatula cDNA clone pkV1-2B16, mRNA  
sequence.  
ACCESSION BE202661  
VERSION BE202661.1 GI:8745925  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 482)

VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,  
Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and  
Fraser, C.M.

ESTs from roots of Medicago truncatula 24 hours after inoculation  
with Sinorhizobium meliloti

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: VandenBosch K  
Department of Biology  
Texas A&M University  
College Station, TX 77843-3258, USA  
Tel: 409 845 7707  
Fax: 409 845 2891

Email: kate@mail.bio.tamu.edu  
Texas A&M University: T262354e  
TIGR sequence name: MTIAD08TK  
More information is available at:  
<http://chrysie.tamu.edu/medicago>

## FEATURES

source  
1. 482  
/organism="Medicago truncatula"  
/cultivar="genotype Al7"  
/db\_xref="taxon:3880"  
/clone\_lib="pkV1-2B16"  
/clone\_lib="KV1"  
/tissue\_type="Seedling roots"  
/dev\_stage="24 hours post-inoculation with Sinorhizobium  
meliloti"  
/lab\_host="E. coli strain XL0LR"  
/note="Vector: pBluescript SK -; Site\_1: EcoRI; Site\_2:  
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
was directionally ligated into the Unizap XR vector from  
Stratagene and packaged using Gigapack III Gold packaging  
extracts. Plasmids containing cDNA inserts were excised  
from the recombinant lambda-zap phage using Ex-assist  
helper phage and propagated in XL0LR cells."  
BASE COUNT 157 a 96 c 106 g 123 t  
ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x BE202661 ..  
Align seg 1/1 to: BE202661 from: 1 to: 482  
1 AspCluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
152 GATGAACCTCTAGCTAGTGGTTACAAAAGTGAATCTTCAGACATGGC 201

17 a 17

202 T 202

seq\_name: gb\_est1:BE205231

## seq\_documentation\_block:

LOCUS BE205231 495 bp mRNA EST 05-SEP-2000  
DEFINITION EST397907 KV0 Medicago truncatula cDNA clone pkV0-21C7, mRNA  
sequence.  
ACCESSION BE205231  
VERSION BE205231.1 GI:8748527  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 495)

VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,  
Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and  
Fraser, C.M.

ESTs from uninoculated seedling roots of Medicago truncatula

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: VandenBosch K  
Department of Biology  
Texas A&M University  
College Station, TX 77843-3258, USA  
Tel: 409 845 7707  
Fax: 409 845 2891

Email: kate@mail.bio.tamu.edu  
Texas A&M University name: T265003e  
TIGR sequence name: MTGBK16TK  
More information is available at:  
<http://chrysie.tamu.edu/medicago>



```

FEATURES
  source
    Location/Qualifiers
      1..544
        /organism="Medicago truncatula"
        /db_xref="taxon:3880"
        /clone="NF018F10NR"
        /clone_lib="Modulated root"
        /tissue_type="root"
        /dev_stage="Pooled developmental"
        /note="Vector: Lambda Zap; Four-week old Rhizobium
        meliloti-inoculated Medicago truncatula roots, containing
        a mixture of young and old roots and nodules."
BASE COUNT      185 a 115 c 107 g 133 t 4 others
ORIGIN

alignment_scores:
  Quality: 68.00      Length: 17
  Ratio: 4.000       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 76.471

alignment_block:
  US-09-485-529-104 x AW684591 ..
    Align seg 1/1 to: AW684591 from: 1 to: 544
      1 AspgLUleuLeuAlaAlaLeuGLyTYrLYsValArgAlaSerAspMetAl 17
      ||||||||||||||||||||||||||||||||||||||||||||||||
      179 GATGAACCTCTAGCTGTAGTTGGTTACAAAGTGAAATCTTCAGACATGCC 228
      17 a 17
      |
      229 T 229
    seq_name: gb_estl:AW693047

  seq_documentation_block:
    LOCUS      AW693047      601 bp      mRNA      EST      21-DEC-2000
    DEFINITION  NF059C03ST1F1020 Developing stem Medicago truncatula cDNA clone
    ACCESSION  AW693047
    VERSION    NF059C03ST 5', mRNA sequence.
    KEYWORDS   EST.
    SOURCE     barrel medic.
    ORGANISM   Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
  REFERENCE   1 (bases 1 to 601)
  AUTHORS     He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
               ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
               ,R.A.
  TITLE       Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula stem library
  JOURNAL     Unpublished (2000)
  COMMENT     On Apr 14, 2000 this sequence version replaced gi:7567783.
               Contact: Dixon RA
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73402, USA
               Tel: 580 221 7302
               Fax: 580 221 7380
               Email: radixon@noble.org
               Insert Length: 653 Std Error: 0.00
               Plate: 059 row: C column: 03
               Seq primer: TCACACAGGAACAGCATGAC.
               Location/Qualifiers
                 1..601
                   /organism="Medicago truncatula"
                   /db_xref="taxon:3880"
                   /clone="NF059C03ST"
                   /clone_lib="Developing stem"
                   /tissue_type="stem"

FEATURES
  source
    Location/Qualifiers
      1..601
        /organism="Medicago truncatula"
        /db_xref="taxon:3880"
        /clone="NF059C03ST"
        /clone_lib="Developing stem"
        /tissue_type="stem"

```

```

/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT      200 a 120 c 128 g 153 t
ORIGIN

alignment_scores:
  Quality: 68.00      Length: 17
  Ratio: 4.000       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 76.471

alignment_block:
  US-09-485-529-104 x AW693047 ..
    Align seg 1/1 to: AW693047 from: 1 to: 601
      1 AspgLUleuLeuAlaAlaLeuGLyTYrLYsValArgAlaSerAspMetAl 17
      ||||||||||||||||||||||||||||||||||||||||||||||||
      175 GATGAACCTCTAGCTGTAGTTGGTTACAAAGTGAAATCTTCAGACATGCC 224
      17 a 17
      |
      225 T 225
    seq_name: gb_estl:AW697328

  seq_documentation_block:
    LOCUS      AW697328      613 bp      mRNA      EST      21-DEC-2000
    DEFINITION  NF115E11ST1F1086 Developing stem Medicago truncatula cDNA clone
    ACCESSION  AW697328
    VERSION    NF115E11ST 5', mRNA sequence.
    KEYWORDS   EST.
    SOURCE     barrel medic.
    ORGANISM   Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
  REFERENCE   1 (bases 1 to 613)
  AUTHORS     He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
               ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
               ,R.A.
  TITLE       Expressed Sequence Tags from the Samuel Roberts Noble Foundation
               Medicago truncatula stem library
  JOURNAL     Unpublished (2000)
  COMMENT     On Apr 14, 2000 this sequence version replaced gi:7572090.
               Contact: Dixon RA
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73402, USA
               Tel: 580 221 7302
               Fax: 580 221 7380
               Email: radixon@noble.org
               Insert Length: 653 Std Error: 0.00
               Plate: 115 row: E column: 11
               Seq primer: TCACACAGGAACAGCATGAC.
               Location/Qualifiers
                 1..613
                   /organism="Medicago truncatula"
                   /db_xref="taxon:3880"
                   /clone="NF115E11ST"
                   /clone_lib="Developing stem"
                   /tissue_type="stem"
                   /dev_stage="Pooled developmental"
                   /note="Vector: Lambda Zap; Contains a mixture of
                   internodal stem segments"
BASE COUNT      202 a 123 c 131 g 157 t
ORIGIN

alignment_scores:

```

Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x AW697328 ..

Align seg 1/1 to: AW697328 from: 1 to: 613

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
183 GATGAACCTCTAGCTAGTGTGTTACAAAGTGAATCTTCAGACATGCG 232

17 a 17

233 T 233

seq\_name: gb\_est1:AW690574

seq\_documentation\_block: 614 bp mRNA EST 15-JUN-2000  
LOCUS AW690574  
DEFINITION NF031C08S1F1000 Developing stem Medicago truncatula cDNA clone  
NFO31C08ST 5', mRNA sequence.  
ACCESSION AW690574  
VERSION AW690574.1 GI:7565310  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 614)  
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon  
,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
JOURNAL Medicago truncatula stem library  
COMMENT Unpublished (2000)

Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7302  
Fax: 580 221 7380  
Email: radixon@noble.org  
Insert Length: 614 Std Error: 0.00  
Plate: 031 row: C column: 08  
Seq primer: TCACACAGGAACAGCTATGAC.

## FEATURES

source  
1..614  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF031C08ST"  
/clone\_lib="Developing stem"  
/tissue\_type="stem"  
/dev\_stage="Pooled developmental"  
/note="Vector: Lambda Zap; Contains a mixture of  
internodal stem segments"

BASE COUNT 183 a 130 c 126 g 170 t 5 others  
ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x AW690574 ..

Align seg 1/1 to: AW690574 from: 1 to: 614

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
38 GATGAACCTCTAGCTAGTGTGTTACAAAGTGAATCTTCAGACATGCG 87

17 a 17

88 T 88

seq\_name: gb\_est1:AW690601

seq\_documentation\_block: 620 bp mRNA EST 20-DEC-2000  
LOCUS AW690601  
DEFINITION NF036C12ST1F1000 Developing stem Medicago truncatula cDNA clone  
NFO36C12ST 5', mRNA sequence.

ACCESSION AW690601

VERSION AW690601.2 GI:11933126

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE 1 (bases 1 to 620)  
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell  
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon  
,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula stem library

JOURNAL Unpublished (2000)

COMMENT On Apr 14, 2000 this sequence version replaced gi:7565337.

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 649 Std Error: 0.00

Plate: 036 row: C column: 12

Seq primer: TCACACAGGAACAGCTATGAC.

## FEATURES

source  
1..620  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF036C12ST"  
/clone\_lib="Developing stem"  
/tissue\_type="stem"  
/dev\_stage="Pooled developmental"  
/note="Vector: Lambda Zap; Contains a mixture of  
internodal stem segments"

BASE COUNT 205 a 122 c 133 g 160 t

ORIGIN

## alignment\_scores:

Quality: 68.00 Length: 17  
Ratio: 4.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 76.471

## alignment\_block:

US-09-485-529-104 x AW690601 ..

Align seg 1/1 to: AW690601 from: 1 to: 620

1 AspGluLeuAlaLeuGlyTyrLysValArgAlaSerAspMetal 17  
|||||  
190 GATGAACCTCTAGCTAGTGTGTTACAAAGTGAATCTTCAGACATGCG 239

17 a 17

240 T 240









/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-544-332-1 +	35.00	85.14	4.3e+03	8457	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-357-251-3 -	33.00	106.98	261.53	451	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-245-041-14 +	35.00	84.99	4.4e+03	8589	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-916-576B-9 -	33.00	106.17	290.11	489	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-245-041-1 +	35.00	84.72	4.5e+03	8827	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-669-408B-3 +	33.00	105.79	304.63	508	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-385-028-13 +	35.00	81.98	6.5e+03	11604	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-391 +	33.00	104.14	376.29	599	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-385-028-1 +	35.00	79.36	9.0e+03	15079	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-798 +	33.00	104.06	380.32	604	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-1 +	35.00	22.60	6.9e+05	4411529	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-989 +	33.00	103.61	403.07	632	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-581 +	34.00	107.74	237.15	607	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-293 +	33.00	102.12	487.43	733	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-328-111-231 +	34.00	107.23	253.30	639	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-989 +	33.00	100.47	602.70	865	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-198-956-3 -	34.00	106.81	267.10	666	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-243-543A-2 -	33.00	99.30	699.89	972	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-347-335A-2 -	34.00	102.50	464.79	1026	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-125-628-7 -	33.00	99.26	703.59	976	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-080-386-1 +	34.00	100.71	584.60	1227	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-125-628-8 -	33.00	99.26	703.59	976	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-390-000A-1 +	34.00	100.71	584.60	1227	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-690-493-2 -	33.00	98.68	758.57	1035	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-850-910A-44 -	34.00	98.57	768.62	1519	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-169-948B-15 +	33.00	97.58	873.11	1155	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-841-178-21 +	34.00	97.06	933.05	1767	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-448-873-15 +	33.00	97.58	873.11	1155	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-496-855A-5 -	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-382-452D-15 +	33.00	97.20	916.95	1200	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-466-589-11 -	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-690-493-3 -	33.00	97.20	916.95	1200	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-700-636-11 -	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-690-493-3 -	33.00	97.20	916.95	1200	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-574-11 -	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-117-121-31 -	33.00	96.50	1.0e+03	1287	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-217-345-11 -	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-353-550-7 -	33.00	95.48	1.1e+03	1425	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-841-178-23 +	34.00	96.26	1.0e+03	1915	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-551-687-7 -	33.00	95.48	1.1e+03	1425	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-841-178-23 +	34.00	96.23	1.0e+03	1921	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-669-408B-9 +	33.00	94.61	1.3e+03	1555	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-785-430-1 +	34.00	96.14	1.1e+03	1938	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-776-210-3 +	33.00	94.32	1.3e+03	1600	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-996-800-1 +	34.00	96.14	1.1e+03	1938	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-075-193-3 +	33.00	92.88	1.6e+03	1848	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-252-126-1 +	34.00	94.29	1.3e+03	2332	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-564-090A-3 +	33.00	92.88	1.6e+03	1848	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-942-761-1 +	34.00	94.29	1.3e+03	2332	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-934-494-1 -	33.00	92.77	1.6e+03	1869	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-07-795-859B-5 -	34.00	93.85	1.4e+03	2437	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-143-068-1 -	33.00	92.77	1.6e+03	1869	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-616-5 -	34.00	93.85	1.4e+03	2437	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-358-683-1 -	33.00	91.50	1.9e+03	2121	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-731-157A-3 +	34.00	93.79	1.4e+03	2451	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-202-389-11 -	33.00	90.80	2.1e+03	2276	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-229-448B-1 +	34.00	93.79	1.4e+03	2451	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-018-129-4 -	33.00	88.76	2.7e+03	2790	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-541-780-3 +	34.00	93.79	1.4e+03	2451	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-448-250-4 -	33.00	88.76	2.7e+03	2790	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-861-464-3 -	34.00	90.36	2.2e+03	3455	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-840-204-1 +	33.00	88.46	2.8e+03	2876	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-396-001-3 -	34.00	90.36	2.2e+03	3455	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-840-204-10 -	33.00	88.46	2.8e+03	2876	!
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-323-433A-3 -	34.00	90.13	2.3e+03	3534	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-537-846A-1 +	33.00	88.08	3.0e+03	2989	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-841-178-24 +	34.00	90.13	2.3e+03	3534	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-664-962B-3 -	33.00	87.94	3.0e+03	3031	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-841-178-25 +	34.00	90.13	2.3e+03	3534	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-110-158-3 -	33.00	87.94	3.0e+03	3031	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-841-178-26 +	34.00	90.13	2.3e+03	3534	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-596-985-1 +	33.00	86.46	3.6e+03	3515	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-936-644-3 +	34.00	89.09	2.6e+03	3924	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-391-743A-1 +	33.00	85.71	4.0e+03	3786	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-3 +	34.00	89.09	2.6e+03	3924	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-295-028-42 +	33.00	85.71	4.0e+03	3786	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-936-644-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-106-582-42 +	33.00	85.71	4.0e+03	3786	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-305-639-4 +	33.00	81.05	7.2e+03	6038	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-790-912-1 +	33.00	80.78	7.5e+03	6201	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-305-639-1 +	33.00	78.72	9.7e+03	7622	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-028-851-1 +	33.00	78.70	9.8e+03	7633	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-815-520-1 +	33.00	78.70	9.8e+03	7633	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-273-163-1 +	33.00	78.70	9.8e+03	7633	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-743-637B-27 +	33.00	76.95	1.2e+04	9100	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-526-840B-27 +	33.00	76.95	1.2e+04	9100	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-926-922-6 +	33.00	69.95	3.0e+04	18318	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-414-926A-6 +	33.00	69.95	3.0e+04	18318	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-253-682-6 +	33.00	69.95	3.0e+04	18318	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-527-657-6 +	33.00	69.95	3.0e+04	18318	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-770-379-19 -	33.00	63.45	6.6e+04	3510	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-757-669A-19 -	33.00	63.45	6.6e+04	3510	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-195-966-1 +	32.50	76.70	1.3e+04	7739	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-299-675-1 +	32.50	76.69	1.3e+04	7745	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-485-241-1 +	32.50	76.69	1.3e+04	7745	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-874-162-1 +	32.50	76.69	1.3e+04	7745	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-874-162-1 +	32.50	76.69	1.3e+04	7745	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-1 +	32.00	120.79	2.4e+04	17372	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-829-876-22 -	32.00	120.79	2.4e+04	17372	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-743-130A-40 +	32.00	110.24	172.10	224	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-347 +	32.00	106.34	283.90	331	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-468-709B-3 +	32.00	101.73	512.84	525	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-241-664B-3 +	32.00	101.73	512.84	525	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-329-704-3 -	32.00	101.05	559.63	562	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-472-604-3 -	32.00	101.05	559.63	562	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-486-117-3 -	32.00	101.05	559.63	562	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-604-3 -	32.00	101.05	559.63	562	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-604-3 -	32.00	101.05	559.63	562	!
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-352-552-1 +	34.00	89.08	2.6e+03	3927	!	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-604-3 -	32.00	101.05	559.63	56	

/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-506-553C-7-	32.00	100.28	617.71	607	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-754-282B-5-	32.00	84.51	4.7e+03	2940
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-235-836C-35+	32.00	100.00	639.97	624	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-671-817A-1+	32.00	84.34	4.8e+03	2990
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-506-553C-22-	32.00	99.58	675.69	651	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-149-100-1+	32.00	84.23	4.8e+03	3024
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-434-255-9+	32.00	99.16	713.16	679	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-671-817A-3+	32.00	84.09	4.9e+03	3066
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-459-967-9+	32.00	99.16	713.16	679	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-453-104-23+	32.00	83.66	5.2e+03	3200
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-460-327-9+	32.00	99.16	713.16	679	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-694-824-23+	32.00	83.66	5.2e+03	3200
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-459-871-6+	32.00	99.16	713.16	679	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-453-104-22+	32.00	83.66	5.2e+03	3201
/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US95-16766-31+	32.00	99.10	718.55	683	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-694-824-22+	32.00	83.66	5.2e+03	3201
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-328-111-695+	32.00	99.00	728.01	690	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-067-676B-2+	32.00	83.30	5.5e+03	3319
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-567-816A-1-	32.00	97.93	835.15	768	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-282-845-1+	32.00	83.30	5.5e+03	3319
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-506-553C-25-	32.00	97.14	923.96	831	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-428-414A-4+	32.00	83.30	5.5e+03	3319
/cgn2_6/pdata2/ina/6A_COMB.seq:US-09-087-194-34+	32.00	96.75	971.26	864	!	/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US94-00334-2+	32.00	83.30	5.5e+03	3333
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-198-955A-9+	32.00	95.21	1.2e+03	1008	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-349-867-18+	32.00	82.87	5.8e+03	3465
/cgn2_6/pdata2/ina/backfiles1.seq:54915260-1+	32.00	94.78	1.3e+03	1052	! Pate	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-239-475-18+	32.00	82.87	5.8e+03	3465
/cgn2_6/pdata2/ina/6A_COMB.seq:US-09-087-088-1+	32.00	94.00	1.4e+03	1137	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-598-305A-18+	32.00	82.87	5.8e+03	3465
/cgn2_6/pdata2/ina/6A_COMB.seq:US-09-074-912-3-	32.00	93.24	1.5e+03	1227	!	/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US95-05431-18+	32.00	82.87	5.8e+03	3465
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-097-938-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-951-715A-1+	32.00	82.86	5.8e+03	3468
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-476-000-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-459-448A-1+	32.00	82.86	5.8e+03	3468
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-472-840-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-459-504B-1+	32.00	82.86	5.8e+03	3468
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-476-976-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-459-444-1+	32.00	82.86	5.8e+03	3468
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-474-410-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-530-492-164+	32.00	82.85	5.8e+03	3471
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-486-673B-3-	32.00	93.02	1.6e+03	1255	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-906-517-164+	32.00	82.85	5.8e+03	3471
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-068-109-1-	32.00	92.88	1.6e+03	1272	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-828-788A-7+	32.00	82.68	5.9e+03	3531
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-312-038-4+	32.00	92.71	1.6e+03	1294	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-920-085-1+	32.00	82.68	5.9e+03	3531
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-476-000-62-	32.00	91.83	1.8e+03	1414	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-920-085-1+	32.00	82.68	5.9e+03	3531
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-472-840-62-	32.00	91.83	1.8e+03	1414	!	/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US92-11397-7+	32.00	82.68	5.9e+03	3531
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-476-976-62-	32.00	91.83	1.8e+03	1414	!	/cgn2_6/pdata2/ina/backfiles1.seq:5169629-1+	32.00	82.68	5.9e+03	3531
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-474-410-62-	32.00	91.83	1.8e+03	1414	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-754-490-27+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-486-673B-62-	32.00	91.83	1.8e+03	1414	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-922-505A-27+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-724-814-15+	32.00	91.66	1.9e+03	1437	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-922-505A-33+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-754-282B-2-	32.00	90.74	2.1e+03	1576	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-260-952A-27+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-906-769-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-253-341-27+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-906-616-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-253-341-33+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-817-795-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-253-331A-27+	32.00	82.67	5.9e+03	3534
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-639-075A-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-377-690-3+	32.00	82.66	5.9e+03	3572
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-012-692-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-713-815A-2+	32.00	82.56	6.0e+03	3572
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-012-692-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-602-737-7+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-906-613-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-754-490-29+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US95-14442A-112+	32.00	90.72	2.1e+03	1580	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-922-505A-29+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/5B_COMB.seq:US-09-318-448-8+	32.00	90.17	2.3e+03	1669	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-001-982-7+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-841-178-20+	32.00	89.73	2.4e+03	1743	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-260-952A-29+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-363-970-6+	32.00	89.48	2.5e+03	1788	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-253-341-29+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-754-282B-3-	32.00	89.26	2.5e+03	1827	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-253-331A-29+	32.00	82.54	6.0e+03	3579
/cgn2_6/pdata2/ina/6A_COMB.seq:US-07-615-509A-1+	32.00	89.21	2.6e+03	1836	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-936-135-5-	32.00	82.34	6.2e+03	3652
/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-615-448A-6+	32.00	89.06	2.6e+03	1864	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-1421-1+	32.00	81.61	6.8e+03	3931
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-196-361-6+	32.00	89.06	2.6e+03	1864	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-735-893-1+	32.00	81.61	6.8e+03	3931
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-446-934-6+	32.00	89.06	2.6e+03	1864	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-446-486-1+	32.00	81.40	6.9e+03	4014
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-448-128-6+	32.00	89.06	2.6e+03	1864	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-463-308-1+	32.00	81.40	6.9e+03	4014
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-948-703-6+	32.00	89.06	2.6e+03	1864	!	/cgn2_6/pdata2/ina/backfiles1.seq:5254799-3+	32.00	80.94	7.3e+03	4201
/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-181-112-17+	32.00	89.02	2.6e+03	1872	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-377-690-2+	32.00	80.61	7.7e+03	4343
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-107-761-3+	32.00	88.52	2.8e+03	1969	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-486-380-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-747-574-1-	32.00	88.46	2.8e+03	1981	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-430-080-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-295-814E-9-	32.00	88.40	2.8e+03	1991	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-250-848-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-343-361-9-	32.00	88.40	2.8e+03	1991	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-438-192-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US93-01959-9-	32.00	88.40	2.8e+03	1991	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-475-971-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/6A_COMB.seq:US-09-023-023-1+	32.00	88.01	3.0e+03	2071	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-09-021-203-1+	32.00	80.57	7.7e+03	4360
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-687-865A-1+	32.00	87.12	3.3e+03	2263	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-754-282B-6-	32.00	80.55	7.7e+03	4370
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-043-711-1+	32.00	87.12	3.3e+03	2263	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-559-303B-72+	32.00	80.40	7.9e+03	4437
/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-312-038-3+	32.00	87.01	3.4e+03	2289	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-175-828-72+	32.00	80.40	7.9e+03	4437
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-618-164-1+	32.00	86.87	3.4e+03	2322	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-09-103-875-4-	32.00	80.34	7.9e+03	4460
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-804-439A-5-	32.00	86.43	3.6e+03	2425	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-631-097-3+	32.00	79.33	9.0e+03	4935
/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-720-229-5-	32.00	86.43	3.6e+03	2425	!	/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-766-014-1+	32.00	78.62	9.9e+03	5300
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-754-282B-4-	32.00	86.37	3.7e+03	2441	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-456-200B-5+	32.00	78.36	1.0e+04	5438
/cgn2_6/pdata2/ina/5B_COMB.seq:US-08-973-675-1+	32.00	85.65	4.0e+03	2623	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-08-810-712-9+	32.00	77.57	1.1e+04	5886
/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-977-434-11+	32.00	85.44	4.1e+03	2679	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-675-566-5-	32.00	77.06	1.2e+04	6196
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-458-819-11+	32.00	85.44	4.1e+03	2679	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-675-566-14-	32.00	76.98	1.2e+04	6243
/cgn2_6/pdata2/ina/PCTUS_COMB.seq:PC1-US91-07035-11+	32.00	85.44	4.1e+03	2679	!	/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-675-566-6-	32.00	76.57	1.3e+04	6503
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-434-255-1+	32.00	85.35	4.2e+03	2702	!	/cgn2_6/pdata2/ina/6A_COMB.seq:US-08-975-762-6+	32.00	75.71	1.4e+04	7091
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-459-967-1+	32.00	85.35	4.2e+03	2702	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-295-028-46+	32.00	75.71	1.4e+04	7091
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-460-327-1+	32.00	85.35	4.2e+03	2702	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-106-582-46+	32.00	75.71	1.4e+04	7091
/cgn2_6/pdata2/ina/5A_COMB.seq:US-08-459-871-1+	32.00	85.35	4.2e+03	2702	!	/cgn2_6/pdata2/ina/6B_COMB.seq:US-09-381-862-4+	32.00	75.64	1.4e+04	7143
/cgn2_6/pdata2/ina/5A_COMB.seq:US-07-671-817A-2+	32.00	84.94	4.4e+03	2815	!					

/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-675-566-13	-	32.00	75.31	1.5e+04	7379	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-232-016-23	+	32.00	75.06	1.6e+04	7566	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-232-016-22	+	32.00	74.96	1.6e+04	7639	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-30	+	31.50	93.74	1.4e+03	969
-	-	32.00	74.14	1.7e+04	8298	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-03076-1	-	32.00	74.07	1.8e+04	8351	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-198-446B-14	-	32.00	74.07	1.8e+04	8351	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-870-693-14	-	32.00	73.76	1.9e+04	8618	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-516-25	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-675-566-21	-	32.00	73.56	1.9e+04	8792	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-675-566-25	-	32.00	73.76	1.9e+04	8792	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-639-857-32	-	32.00	73.17	2.0e+04	9143	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-759-038-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-643-734C-1	+	32.00	71.12	2.6e+04	11219	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-30	+	31.50	93.74	1.4e+03	969
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-439-009A-1	+	32.00	71.12	2.6e+04	11219	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-848-932-3	-	31.50	93.69	1.4e+03	974
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-646-695-1	-	32.00	68.69	3.5e+04	14311	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-008-180-3	-	31.50	93.69	1.4e+03	974
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-646-695-7	-	32.00	68.69	3.5e+04	14311	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-073-384C-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06053-7	+	32.00	68.69	3.5e+04	14311	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06053-7	+	32.00	68.69	3.5e+04	14311	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-125-468-1	+	32.00	61.29	8.6e+04	30001	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-238-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-474-933-1	+	32.00	61.29	8.6e+04	30001	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-733-609-4	-	32.00	59.95	1.0e+05	34303	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-733-609-4	-	32.00	59.95	1.0e+05	34303	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-315-372-4	-	32.00	59.95	1.0e+05	34303	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-244-752-4	-	32.00	59.95	1.0e+05	34303	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-245-497-4	-	32.00	59.95	1.0e+05	34303	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-374-483-6	-	32.00	59.93	1.0e+05	34382	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-973-334-3	-	32.00	59.64	1.0e+05	35408	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-563-869A-3	-	32.00	59.64	1.0e+05	35408	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-12	+	31.50	88.72	2.7e+03	1600
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-549-489-3	-	32.00	59.64	1.0e+05	35408	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-622-166A-1	-	31.50	88.67	2.7e+03	1608
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-735-609-1	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-073-384C-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-735-609-1	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-379-452-43	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-315-372-1	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-238-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-244-752-1	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-245-497-1	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-409-670-43	-	32.00	59.49	1.1e+05	35935	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-338-907-183	-	32.00	58.94	1.1e+05	37950	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-1	+	32.00	57.63	1.3e+05	43280	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-996-306-1	-	32.00	54.96	1.8e+05	56516	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-338-907-1	-	32.00	54.96	1.8e+05	56516	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-338-907-179	-	32.00	54.96	1.8e+05	56520	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-9	+	31.50	88.43	2.8e+03	1647
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-708-294-4	-	32.00	51.46	2.6e+05	80246	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-848-932-1	+	31.50	88.32	2.9e+03	1666
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-078-294-3	-	32.00	51.41	2.6e+05	80595	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-008-180-1	+	31.50	88.32	2.9e+03	1666
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-245-041-5	-	32.00	50.31	2.9e+05	90050	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-073-384C-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-245-041-5	-	32.00	50.31	2.9e+05	90050	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-31	+	31.50	102.51	464.18	403	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-997-080-163	-	31.50	102.51	464.18	403	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-238-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-997-362-163	-	31.50	102.51	464.18	403	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-066B-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-095-855-163	-	31.50	96.39	1.0e+03	743	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-095-855-198	-	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-10	+	31.50	86.06	3.8e+03	2088
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-899-336-2	+	31.50	85.11	4.3e+03	2296
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-899-336-2	+	31.50	85.11	4.3e+03	2296
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-31	+	31.50	93.96	1.4e+03	948	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-977-434-1	+	31.50	84.27	4.8e+03	2499
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-073-384C-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-819-1	+	31.50	84.27	4.8e+03	2499
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-07-602-848E-1	+	31.50	84.27	4.8e+03	2499
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-587-856-1	+	31.50	84.27	4.8e+03	2499
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-238-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US91-07035-1	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-973-384C-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-238-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-11	+	31.50	93.81	1.4e+03	962	1	/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-21	+	31.50	84.26	4.8e+03	2502
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-7												



```

/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-167 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-187 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-189 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-65 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-68 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-70 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-106 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-129 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-131 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-106 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-129 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-131 + 31.50 84.24 4.8e+03 2505
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-073-334C-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-254-359A-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-043-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-481-238-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-484-956-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-757-653-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-491-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-756-386-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-823-516-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-682-853A-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-759-038-1 + 31.50 84.24 4.8e+03 2506
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-758-314-1 + 31.50 84.24 4.8e+03 2507
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-066B-1 + 31.50 84.24 4.8e+03 2507
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-156-020-1 + 31.50 83.77 5.1e+03 2626
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-156-020-3 + 31.50 83.77 5.1e+03 2626
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-156-020-5 + 31.50 83.77 5.1e+03 2626
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-156-020-7 + 31.50 83.77 5.1e+03 2626
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-156-020-9 + 31.50 83.77 5.1e+03 2626
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-105-697-9 + 31.50 83.56 5.3e+03 2682
/cgn2_6/ptodata/2/ina/backfill1.seq:5244792-12 - 31.00 110.93 157.64 144
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-687-080-79 + 31.00 110.00 177.55 158
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-307-489-42 + 31.00 106.33 284.13 228
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-299-268-42 + 31.00 106.33 284.13 228
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-480 + 31.00 105.17 339.61 256
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-297 + 31.00 105.06 334.57 259
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-060-756-628 + 31.00 98.13 813.56 518
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-036-355A-1 + 31.00 97.57 874.45 548
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-441-507-17 + 31.00 96.76 969.64 594
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-685-466C-3 + 31.00 96.46 1.0e+03 612
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-752 + 31.00 96.46 1.0e+03 612
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-661-610C-14 + 31.00 96.27 1.0e+03 624
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-1029 + 31.00 95.24 1.2e+03 692
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-613 - 31.00 94.91 1.2e+03 715
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-639 + 31.00 94.70 1.3e+03 730
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-248-335-39 + 31.00 94.13 1.4e+03 773
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-436-748-2 + 31.00 93.66 1.4e+03 810
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-483-857-2 + 31.00 93.66 1.4e+03 810
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-349-637-2 + 31.00 93.55 1.5e+03 819
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-391-916A-7 + 31.00 93.48 1.5e+03 825
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-391-916A-5 + 31.00 91.87 1.8e+03 969
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-885-970A-7 - 31.00 91.67 1.9e+03 989
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-298-687A-7 - 31.00 91.67 1.9e+03 989
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-530-829-7 - 31.00 91.67 1.9e+03 989
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-298-839-7 - 31.00 91.67 1.9e+03 989
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-371-913A-10 + 31.00 91.15 2.0e+03 1041
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-859-167-3 - 31.00 91.12 2.0e+03 1045
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-109-273-3 - 31.00 91.12 2.0e+03 1045
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-415-751-47 - 31.00 90.73 2.1e+03 1086
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-248-335-53 + 31.00 90.60 2.1e+03 1100
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-911-853-26 + 31.00 90.19 2.3e+03 1146
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-479-409-26 + 31.00 90.19 2.3e+03 1146
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-383-750-5 - 31.00 90.06 2.3e+03 1161
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-684-687-1 - 31.00 90.06 2.3e+03 1161
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-352-678-5 - 31.00 90.06 2.3e+03 1161
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-391-916A-3 + 31.00 89.86 2.4e+03 1164
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-391-916A-5 - 31.00 89.76 2.4e+03 1195
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-242-663A-1 + 31.00 89.61 2.4e+03 1215
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-954-536-17 + 31.00 89.61 2.4e+03 1215

```

```

/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-748-547-1 + 31.00 89.61 2.4e+03 1215
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-573-890-1 + 31.00 89.60 2.4e+03 1216
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-248-179-1 + 31.00 89.60 2.4e+03 1216
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-477-254B-3 - 31.00 89.41 2.5e+03 1239
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-734B-3 - 31.00 89.41 2.5e+03 1239
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-428-734B-3 - 31.00 89.41 2.5e+03 1239
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-713-556F-3 - 31.00 89.41 2.5e+03 1239
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-391-916A-1 - 31.00 89.34 2.5e+03 1248
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-672-56A-10 - 31.00 89.32 2.5e+03 1251
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-672-56A-10 - 31.00 89.27 2.5e+03 1257
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-961-083-115 + 31.00 89.19 2.6e+03 1267
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-289-349-9 + 31.00 89.01 2.6e+03 1290
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-370-190-2 + 31.00 88.89 2.7e+03 1306
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-537-400-2 + 31.00 88.86 2.7e+03 1310
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-712-709-2 - 31.00 88.58 2.8e+03 1347
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-111-442-2 - 31.00 88.58 2.8e+03 1347
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-541-228-2 - 31.00 88.58 2.8e+03 1347
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-018-584A-43 + 31.00 88.44 2.8e+03 1366
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-394-189B-3 + 31.00 88.39 2.8e+03 1373
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-657-392-26 - 31.00 88.39 2.8e+03 1373
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-657-392-26 - 31.00 88.24 2.9e+03 1394
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-255-671-3 + 31.00 88.24 2.9e+03 1394
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-239-909-1 + 31.00 88.19 2.9e+03 1401
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-56 + 31.00 88.18 2.9e+03 1402
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-54 + 31.00 88.18 2.9e+03 1402
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-657-392-1 - 31.00 87.81 3.1e+03 1454
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-657-392-1 - 31.00 87.81 3.1e+03 1454
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-657-392-26 - 31.00 87.81 3.1e+03 1455
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-824-707-1 + 31.00 87.79 3.1e+03 1458
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-824-707-1 + 31.00 87.78 3.1e+03 1459
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-225-477B-2 + 31.00 87.38 3.2e+03 1519
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-867-057-2 + 31.00 87.21 3.3e+03 1544
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-128-369-2 + 31.00 87.21 3.3e+03 1544
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-466-337A-17 + 31.00 87.18 3.3e+03 1550
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-475-359-17 + 31.00 87.18 3.3e+03 1550
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-802-322-2 + 31.00 87.18 3.3e+03 1550
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-465-887A-17 + 31.00 87.18 3.3e+03 1550
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-895-601-3 + 31.00 87.18 3.3e+03 1550
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-477-254A-6 - 31.00 86.94 3.4e+03 1587
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-576B-6 - 31.00 86.94 3.4e+03 1587
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-428-734B-6 - 31.00 86.94 3.4e+03 1587
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-713-556F-6 - 31.00 86.94 3.4e+03 1587
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-477-254A-1 - 31.00 86.56 3.6e+03 1649
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-472-576B-1 - 31.00 86.56 3.6e+03 1649
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-428-734B-1 - 31.00 86.56 3.6e+03 1649
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-713-556F-1 - 31.00 86.56 3.6e+03 1649
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-663-566A-10 - 31.00 86.05 3.8e+03 1734
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-023-610-10 - 31.00 86.05 3.8e+03 1734
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-288-065A-10 - 31.00 86.05 3.8e+03 1734
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-362-240A-10 - 31.00 86.05 3.8e+03 1734
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-804-372A-8 - 31.00 86.05 3.8e+03 1734
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-472-028A-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-808-931-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-808-323-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-050-603A-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-102-420B-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-071-296-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-196-268-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-015-683-3 + 31.00 86.03 3.8e+03 1738
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-191-998-3 - 31.00 86.02 3.8e+03 1740
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-672-56A-2 - 31.00 86.02 3.8e+03 1740
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-226-741-2 + 31.00 85.77 4.0e+03 1784
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-436-748-1 + 31.00 85.68 4.0e+03 1800
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-857-1 + 31.00 85.68 4.0e+03 1800
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-820-154A-29 - 31.00 85.10 4.3e+03 1907
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-097-554A-29 - 31.00 85.10 4.3e+03 1907
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-480-640A-29 - 31.00 85.10 4.3e+03 1907
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-295-802-29 - 31.00 85.10 4.3e+03 1907
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-488-237A-29 - 31.00 85.10 4.3e+03 1907

```



[illegible]



```

seq_documentation_block:
; Sequence 4, Application US/08981256A
; Patent No. 6046042
; GENERAL INFORMATION:
; APPLICANT: Meinhard HASSLACHER et al.
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

```

ATTORNEY/AGENT INFORMATION:

```

; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-324 (WSU)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (313) 689-3554
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2005 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..1250
US-07-916-901-5

```

alignment_scores:		
Quality:	42.00	Length: 13
Ratio:	3.818	Gaps: 0
Percent Similarity:	84.615	Percent Identity: 76.923

```
alignment_block:
US-09-485-529-104 x US-07-916-901-5/rev
..
```

Align seg 1/1 to reverse of: US-07-916-901-5 from: 1 to: 2005

3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15  
:::|||||  
130 GTGTTGGCTGCACCTGGGGTCCAAGGTGGGGCGCTGTGAC 92

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-07-991-867B-43

```

seq_documentation_block:
; Sequence 43, Application US/07991867B
; Patent No. 5476781
;
; GENERAL INFORMATION:
;
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: NO. 5476781el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991.867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C3
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-991-867B-43

alignment_scores:
    Quality: 41.00      Length: 14
    Ratio: 3.727       Gaps: 0
    Percent Similarity: 78.571    Percent Identity: 64.286

alignment_block:
    US-09-485-529-104 x US-07-991-867B-43      ..

Align seg 1/1 to: US-07-991-867B-43 from: 1 to: 235

      2 GuUeUeUaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ||| ||| ||| :: ||| ||| ||| ||| ||| |||
      84 GAAATTATTGAACGTCNAAAGTTATCTCTGTAAGAATCCGAT 125

seq_name: /cqn2_6/ptodata/2/ina/5A_COMB.seq:US-07-991-867B-45

```

```

1  seq_documentation_block:
2  ; Sequence 45, Application US/07991867B
3  ; Patent No. 5476781
4  ;
5  GENERAL INFORMATION:
6  ;
7  ; APPLICANT: Moyer, Richard W.
8  ; APPLICANT: Hall, Richard L.
9  ; APPLICANT: Gruidl, Michael E.
10 ; TITLE OF INVENTION: NO. 5476781el Entomopoxvirus Expression System
11 ; NUMBER OF SEQUENCES: 66
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: David R. Saliwanchik
14 ; STREET: 2421 N.W. 41st Street, Suite A-1
15 ; CITY: Gainesville
16 ; STATE: FL
17 ; COUNTRY: USA
18 ; ZIP: 32606
19 ;
20 COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: Floppy disk
22 ; COMPUTER: IBM PC compatible
23 ; OPERATING SYSTEM: PC-DOS/MS-DOS
24 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
25 ;
26 CURRENT APPLICATION DATA:
27 ; APPLICATION NUMBER: US/07/991.867B
28 ; FILING DATE: 12-DEC-1992
29 ; CLASSIFICATION: 435
30 ;
31 PRIOR APPLICATION DATA:
32 ; APPLICATION NUMBER: WO 92/14818
33 ; FILING DATE: 12-FEB-1992
34 ;
35 PRIOR APPLICATION DATA:
36 ; APPLICATION NUMBER: US 07/827,685
37 ; FILING DATE: 30-JAN-1992
38 ;
39 PRIOR APPLICATION DATA:
40 ; APPLICATION NUMBER: US 07/657,584
41 ; FILING DATE: 19-FEB-1991
42 ;
43 ATTORNEY/AGENT INFORMATION:
44 ; NAME: Saliwanchik, David R.
45 ; REGISTRATION NUMBER: 31,794
46 ; REFERENCE/DOCKET NUMBER: UFI14.C3
47 ;
48 TELECOMMUNICATION INFORMATION:
49 ; TELEPHONE: 904-375-8100
50 ; TELEFAX: 904-372-5800
51 ;
52 INFORMATION FOR SEQ ID NO: 45:
53 ; SEQUENCE CHARACTERISTICS:
54 ; LENGTH: 235 base pairs
55 ; TYPE: nucleic acid

```

LENGTH: 200 base pairs  
TYPE: nucleic acid

LENGTH: 200 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-544-332-45

alignment\_scores:  
Quality: 41.00 Length: 14  
Ratio: 3.727 Gaps: 0  
Percent Similarity: 78.571 Percent Identity: 64.286

alignment\_block:  
US-09-485-529-104 x US-08-544-332-45 ..

Align seg 1/1 to: US-08-544-332-45 from: 1 to: 235

2 GluLeuLeuAlaLaLeuGlyTyrLysValArgAlaSerAsp 15  
||||||| ::| ||||| ||||| ||||| |||||  
84 GAATTATTGAACGTCAAAGGTTATCTCTGTTAAAGCATCCGAT 125

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-424-224-1

seq\_documentation\_block:  
; Sequence 1, Application US/08424224  
; Patent No. 5912173  
; GENERAL INFORMATION:  
; APPLICANT: LEONARD, WARREN J.  
; TITLE OF INVENTION: MURINE IL-2R CDNA AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVE.  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT # 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/424,224  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/121,435  
; FILING DATE: 14-SEPT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILLIAM S. FEILER  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4061US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1608  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
; MOLECULE TYPE:  
; DESCRIPTION: OLIGONUCLEOTIDE  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: MURINE  
; INDIVIDUAL ISOLATE: IL-2R  
US-08-424-224-1

alignment\_scores:  
Quality: 41.00 Length: 13  
Ratio: 3.727 Gaps: 0  
Percent Similarity: 84.615 Percent Identity: 69.231

alignment\_block:  
US-09-485-529-104 x US-08-424-224-1/rev ..

Align seg 1/1 to reverse of: US-08-424-224-1 from: 1 to: 1608

2 GluLeuLeuAlaLaLeuGlyTyrLysValArgAlaSer 14  
||||||| ||||| ||||| ||||| ||||| |||||  
1409 GAGCTCCAGCACCCCTGGGCTACAGGTAAGACTCTCT 1371

seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:PCT-US94-02891-68

seq\_documentation\_block:  
; Sequence 68, Application PC/TUS9402891  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN  
; APPLICANT: SERVICES  
; APPLICANT: OFFICE OF TECHNOLOGY TRANSFER, NATIONAL  
; APPLICANT: INSTITUTES OF HEALTH, BOX OTT, BETHESDA, MARYLAND 20892 USA  
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: XSCID  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVE.  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT # 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/02891  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/031,143  
; FILING DATE: 12-MAR-1993  
; APPLICATION NUMBER: 08/121,435  
; FILING DATE: 14-SEPT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILLIAM S. FEILER  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4061  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1608  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
; MOLECULE TYPE: OLIGONUCLEOTIDE  
; DESCRIPTION: NO  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: MURINE  
; INDIVIDUAL ISOLATE: IL-2R  
PCT-US94-02891-68

alignment\_scores:



```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-276-531-31
seq_documentation_block:
; Sequence 31, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murty, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT01
; CLONE: 713784
; US-09-276-531-31

alignment_scores:
; Quality: 39.00 Length: 10
; Ratio: 3.900 Gaps: 0
; Percent Similarity: 100.000 Percent Identity: 70.000

seq_documentation_block:
; US-09-485-529-104 x US-09-276-531-31
; Align seg 1/1 to: US-09-276-531-31 from: 1 to: 703
;
; 1 AspGluLeuLeuAlaLeuGlyTyrIys 10
; :|||||:
; 424 GAGCAGCTGCTGGCGCCTGGCTACGC 453
;
; seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-087-772A-3
; seq_documentation_block:
; Sequence 3, Application US/0808772A
```

```
; Patent No. 5691155
; GENERAL INFORMATION:
; APPLICANT: Nahmias, Clara
; APPLICANT: Emorine, Jean L.
; APPLICANT: Strosberg, Donny A.
; TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine
; TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5691155th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,772A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Linker, Raymond O.
; REGISTRATION NUMBER: 26,419
; REFERENCE/DOCKET NUMBER: 3339-195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-087-772A-3

alignment_scores:
; Quality: 39.00 Length: 13
; Ratio: 3.545 Gaps: 0
; Percent Similarity: 84.615 Percent Identity: 69.231

seq_documentation_block:
; US-09-485-529-104 x US-08-087-772A-3/rev
; Align seg 1/1 to reverse of: US-08-087-772A-3 from: 1 to: 1164
;
; 3 LeuLeuAlaAlaLeuGlyTyrIysValargAlaSerAsp 15
; :|||||:
; 80 GTGTGGCTGCTGGGTCCAGGTAGGCGGCCGAC 42
;
; seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-087-772A-4
; seq_documentation_block:
; Sequence 4, Application US/0808772A
; Patent No. 5691155
; GENERAL INFORMATION:
; APPLICANT: Nahmias, Clara
; APPLICANT: Emorine, Jean L.
; APPLICANT: Strosberg, Donny A.
; TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine
; TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5691155th Carolina
; COUNTRY: USA
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 568..1731
; US-08-087-772A-1

alignment_scores:
    Quality: 39.00      Length: 13
    Ratio: 3.545      Gaps: 0
    Percent Similarity: 84.615      Percent Identity: 69.231

alignment_block:
    US-09-485-529-104 x US-08-087-772A-1/rev ..
    Align seg 1/1 to reverse of: US-08-087-772A-1 from: 1 to: 1920
    3 LeuLeuAlaaleuclyTyrlsValargAlaSerasp 15
    :::::::::::::::::::::||||| |
    647 GTGTTGGCTGCAGTGGGTCACGGTAGGGCGGCCGAC 609

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-450-962-3

seq_documentation_block:
; Sequence 3, Application US/08450962
; Patent No. 6274706
; GENERAL INFORMATION:
; APPLICANT: EMORINE, Laurent; MARULLO, Stefano;
; APPLICANT: STROSBURG, Donny
; TITLE OF INVENTION: INTRON/EXON OF THE HUMAN AND
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KECK, MAHIN & CATE
; STREET: P.O. BOX 06110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606-0110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/2" diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08450,962
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,829
; FILING DATE: 08-SEPT-1993
; APPLICATION NUMBER: 07/721,571
; FILING DATE: 25-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR89/00918
; FILING DATE: 25-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleit, Martin; Gollin, Michael A.
; REGISTRATION NUMBER: 16,900; 31,957
; REFERENCE/DOCKET NUMBER: 47078-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 3:

```

SEQUENCE CHARACTERISTICS:

LENGTH: 3437 bases  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-08-450-962-3

alignment\_scores:  
 Quality: 39.00 Length: 13  
 Ratio: 3.545 Gaps: 0  
 Percent Similarity: 84.615 Percent Identity: 69.231

alignment\_block:

US-09-485-529-104 x US-08-450-962-3/rev ..

Align seg 1/1 to reverse of: US-08-450-962-3 from: 1 to: 3437

3 LeuLeuAlaLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15

:::|||||:::|||||:::|||||:::|||||:::|||||

647 GTGTTGGCTGCTGGGTCCAGGTAGGCGTCCGAC 609

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-137-614A-3

seq\_documentation\_block:

Sequence 3, Application US/08137614A

Patent No. 5487976

GENERAL INFORMATION:

APPLICANT: Soderlund, David M.

APPLICANT: Knipple, Douglas C.

APPLICANT: Henderson, Joseph E.

TITLE OF INVENTION: Gene Encoding An Insect

CORRESPONDENCE ADDRESS: Gamma-Aminobutyric Acid (GABA) Receptor Subunit

NUMBER OF SEQUENCES: 31

ADDRESS: Nixon, Hargrave, Devans & Doyle

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/137,614A

FILING DATE: 15-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Timian, Susan J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 19603/120

TELEPHONE: (716)263-1636

TELEFAX: (716)263-1600

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1491 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-137-614A-3

alignment\_scores:

Quality: 38.00 Length: 16  
 Ratio: 3.455 Gaps: 0  
 Percent Similarity: 68.750 Percent Identity: 43.750

alignment\_block:

US-09-485-529-104 x US-08-137-614A-3 ..

Align seg 1/1 to: US-08-137-614A-3 from: 1 to: 1491

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMet 16

:::|||||:::|||||:::|||||:::|||||:::|||||

374 GAGACTTTGCTGAAAGATATGGGTACCGGATACGTTCTTCGCCAATG 421

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-768-301-1

seq\_documentation\_block:

Sequence 1, Application US/08768301

Patent No. 5854002

GENERAL INFORMATION:

APPLICANT: Tomalski, Michael D.

APPLICANT: Gant, Daniel B.

TITLE OF INVENTION: METHOD OF IDENTIFYING COMPOUNDS THAT BIND

TO THE INSECT GABA RECEPTOR

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: U.S.A.

ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/768,301

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MacLeod, Janet M.

REGISTRATION NUMBER: 35,263

REFERENCE/DOCKET NUMBER: A30693

TELEPHONE: 212-408-2500

TELEFAX: 212-765-2519

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1491 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-768-301-1

alignment\_scores:

Quality: 38.00 Length: 16  
 Ratio: 3.455 Gaps: 0  
 Percent Similarity: 68.750 Percent Identity: 43.750

alignment\_block:

US-09-485-529-104 x US-08-768-301-1 ..

Align seg 1/1 to: US-08-768-301-1 from: 1 to: 1491

1 AspGluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMet 16

:::|||||:::|||||:::|||||:::|||||:::|||||

374 GAGACTTTGCTGAAAGATATGGGTACCGGATACGTTCTTCGCCAATG 421



```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-799-464A-14
seq_documentation_block:
; Sequence 14, Application US/08799464A
; Patent No. 5998601
; GENERAL INFORMATION:
; APPLICANT: Murtaugh, Michael P. et al.
; TITLE OF INVENTION: VR-2332 VIRAL NUCLEOTIDE SEQUENCE AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John M. Collins
; STREET: 2405 Grand Blvd., Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: USA
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,464A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,941
; FILING DATE: August 5, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 22907
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (816) 474-9050
; TELEFAX: (816) 474-9057
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15101 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arteriviridae
; STRAIN: VR-2332
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7384..11775
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /label= ORF1b
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11786..12535
; OTHER INFORMATION: /standard_name= "LV ORF 2"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 212..7402
; OTHER INFORMATION: /standard_name= "LV ORF 1a"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12394..13191
; OTHER INFORMATION: /standard_name= "LV ORF 3"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: 12936..13487
; OTHER INFORMATION: /standard_name= "LV ORF 4"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 13484..14089
; OTHER INFORMATION: /standard_name= "LV ORF 5"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14077..14598
; OTHER INFORMATION: /standard_name= "LV ORF 6"
; OTHER INFORMATION: /citation= ({1})
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14588..14974
; OTHER INFORMATION: /standard_name= "LV ORF 7"
; OTHER INFORMATION: /citation= ({1})
; PUBLICATION INFORMATION:
; AUTHORS: Meulenbergh, J. J. M.
; AUTHORS: Hulst, M. M.
; AUTHORS: de Veijer, E. J.
; AUTHORS: Moonen, P. L.
; AUTHORS: den Besten, A.
; AUTHORS: de Kluuyver, E. P.
; AUTHORS: Wensvoort, G.
; AUTHORS: Moormann, R. J.
; TITLE: Lelystad virus, the causative agent of
; TITLE: porcine epidemic abortion and respiratory
; TITLE: syndrome (PEARS) is related to LDV and EAV.
; JOURNAL: Virology
; VOLUME: 192
; PAGES: 62-72
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 14: FROM 1 TO 15101
; US-08-799-464A-14

alignment_scores:
Quality: 38.00 Length: 15
Ratio: 2.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 46.667

alignment_block:
US-09-485-529-104 x US-08-799-464A-14 ..
Align seg 1/1 to: US-08-799-464A-14 from: 1 to: 15101

1 AspGluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9088 GACCGCATCTGCTGCTCTTGCATATCACATGAAGCGCAGAAC 9132

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-09927-14
seq_documentation_block:
; Sequence 14, Application PC/TUS9509927
; GENERAL INFORMATION:
; APPLICANT: Murtaugh, Michael P.
; TITLE OF INVENTION: VR-2332 VIRAL NUCLEOTIDE SEQUENCE AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John M. Collins
; STREET: 1101 Walnut, Suite 1400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: USA
; ZIP: 64106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
```

```

;
; AUTHORS: Moormann, R. J.
; TITLE: Lelystad virus, the causative agent of
; TITLE: porcine epidemic abortion and respiratory
; TITLE: syndrome (PEARS) is related to LDV and EAV.
; JOURNAL: Virology
; VOLUME: 192
; PAGES: 62-72
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 14: FROM 1 TO 15101
PCT-US95-09927-14

alignment_scores:
    Quality: 38.00      Length: 15
    Ratio: 2.714        Gaps: 0
    Percent Similarity: 93.333   Percent Identity: 46.667

alignment_block:
US-09-485-529-104 x PCT-US95-09927-14 ..
Align seg 1/1 to: PCT-US95-09927-14 from: 1 to: 15101
1 AspLLeuLleuAlaLeuGlyTyrLysValArgAlaserAsp 15
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9088 GACCGCATCTGGCTGTCTGTGCATATCATGAAGCGCAGAAC 9132

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-157-005-1
seq_documentation_block:
; Sequence 1, Application US/08157005
; Patent No. 5620691
; GENERAL INFORMATION:
; APPLICANT: Wensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; APPLICANT: Meulenbergh, Johanna J
; TITLE OF INVENTION: CAUSATIVE AGENT OF THE MYSTERY SWINE DISEASE,
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,005
; FILING DATE: 26-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00096
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, Thomas F
; REGISTRATION NUMBER: 16,579
; REFERENCE/DOCKET NUMBER: 44819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 315-1931
; TELEX: 422523 COOP UI

```

```

CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS95
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,863
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/157,005
FILING DATE: 26-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91201398.4
FILING DATE: 06-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 92200781.0
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00096
FILING DATE: 05-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Turner, Allen C.
REGISTRATION NUMBER: 33041
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (801) 532-1922
TELEFAX: (801) 531-9188
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 15108 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 212..7399
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 7384..11772
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 11786..12532
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 12394..13188
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 12936..13484
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 13484..14086
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 14077..14595
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: 14588..14971
OTHER INFORMATION:
US-08-747-863-1

```

```

4 LeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
||||| |||||||: ||:::|||||
23802 CTCGCGCATGGCTACCGGCTAGAGGCTCTGAC 23837

seq_name: /cgn2_6/ptodat2/2/ina/6B_COMB.seq:US-08-998-416-955

seq_documentation_block:
; Sequence 955, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 955:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 640 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1590RP
; US-08-998-416-955

alignment_scores:
Quality: 37.00 Length: 15
Ratio: 3.364 Gaps: 0
Percent Similarity: 73.333 Percent Identity: 46.667

alignment_block:
US-09-485-529-104 x US-08-998-416-955 ..

Align seg 1/1 to: US-08-998-416-955 from: 1 to: 640

3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetal 17
||||| |||||||: ||:::|||||
362 CTGCTAAACTACTCGGTCATCGAATCGGATTTCGCTACTTTAGCC 406

```

1

```

; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NO. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGCI976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 386:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1292UP
; US-08-998-416-386

alignment_scores:
Quality: 37.00 Length: 15
Ratio: 3.364 Gaps: 0
Percent Similarity: 73.333 Percent Identity: 46.667

alignment_block:
US-09-485-529-104 x US-08-998-416-386/rev ..
Align seg 1/1 to reverse of: US-08-998-416-386 from: 1 to: 798
3 LeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
||||| |||||||:|||||: ||| :|||
427 CTGCTAAACTACTCGGGTATCGAATCCGATTTCTACTTTAGCC 383

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-998-416-497

seq_documentation_block:
; Sequence 497, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park

```

```
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 497:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: PAG1351RP
; US-08-998-416-497

alignment_scores:
  Quality: 37.00      Length: 15
  Ratio: 3.364       Gaps: 0
  Percent Similarity: 73.333   Percent Identity: 46.667

alignment_block:
US-09-485-529-104 x US-08-998-416-497 ..
Align seg 1/1 to: US-08-998-416-497 from: 1 to: 832
3 LeuLeuAlaLeuGlyTyrLysValArgAlaSerAspMetAla 17
||||| ||||||| ||||||| ||||||| |||||||
362 CTGCTAAACTACTCGGGTATCGAATCCGATTGCTTACTTTAGCC 406

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-07-809-457A-14

seq_documentation_block:
; Sequence 14, Application US/07809457A
; Patent No. 5512466
; GENERAL INFORMATION:
; APPLICANT: Klee, Harry J.
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: Control of Fruit Ripening and Senescence
; TITLE OF INVENTION: in Plants
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/809,457A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
```

```
; FILING DATE: 19911217
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,440
; FILING DATE: 26-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10538)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1377 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-809-457A-14

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
  Percent Similarity: 85.714   Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-07-809-457A-14 ..
Align seg 1/1 to: US-07-809-457A-14 from: 1 to: 1377
2 GluLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15
||||| ||||||| ||||||| ||||||| |||||||
1314 GATTTCATGGCTGCTTGGAGCTAAGATCGAAGTCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-476-008-9

seq_documentation_block:
; Sequence 9, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
```

```
/ FILING DATE: 31-AUG-1990
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hoerner Jr., Dennis R.
/ REGISTRATION NUMBER: 30,914
/ REFERENCE/DOCKET NUMBER: 38-21(10660)A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (314)537-6099
/ TELEFAX: (314)537-6047
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1377 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-476-008-9

alignment_scores:
  Quality: 37.00 Length: 14
  Ratio: 3.083 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-08-476-008-9 ..
Align seg 1/1 to: US-08-476-008-9 from: 1 to: 1377

2 GluLeuLeuAlaLaLeuGlyTyrlsValArgAlaSerAsp 15
:::||||:||||:||||:||||:||||:
1314 GATTGTGCTGCTGTGGAGCTAAGATCGAAGCTCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-306-063-9

seq_documentation_block:
/ Sequence 9, Application US/08306063
/ Patent No. 5633435
/ GENERAL INFORMATION:
/ APPLICANT: Barry, Gerard F.
/ APPLICANT: Kishore, Ganesh M.
/ APPLICANT: Padgett, Stephen R.
/ APPLICANT: Stallings, William C.
/ TITLE OF INVENTION: Glyphosate Tolerant
/ TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
/ NUMBER OF SEQUENCES: 69
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
/ STREET: 700 Chesterfield Village Parkway
/ CITY: St. Louis
/ STATE: Missouri
/ COUNTRY: USA
/ ZIP: 63198
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/306,063
/ FILING DATE: 13-SEP-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/749,611
/ FILING DATE: 28-AUG-1991
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/576,537
/ FILING DATE: 31-AUG-1990
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hoerner Jr., Dennis R.
/ REGISTRATION NUMBER: 30,914
```

```
/ REFERENCE/DOCKET NUMBER: 38-21(10660)A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (314)537-6099
/ TELEFAX: (314)537-6047
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1377 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-306-063-9

alignment_scores:
  Quality: 37.00 Length: 14
  Ratio: 3.083 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-08-306-063-9 ..
Align seg 1/1 to: US-08-306-063-9 from: 1 to: 1377

2 GluLeuLeuAlaLaLeuGlyTyrlsValArgAlaSerAsp 15
:::||||:||||:||||:||||:||||:
1314 GATTGTGCTGCTGTGGAGCTAAGATCGAAGCTCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-553-943-14

seq_documentation_block:
/ Sequence 14, Application US/08553943
/ Patent No. 5702933
/ GENERAL INFORMATION:
/ APPLICANT: Klee, Harry J.
/ APPLICANT: Kishore, Ganesh M.
/ TITLE OF INVENTION: Control of Fruit Ripening and Senescence
/ TITLE OF INVENTION: in Plants
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Monsanto Co. BB4F
/ STREET: 700 Chesterfield Village Parkway
/ CITY: St. Louis
/ STATE: Missouri
/ COUNTRY: USA
/ ZIP: 63198
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/553,943
/ FILING DATE: 06-NOV-1995
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/809,457
/ FILING DATE: 17-DEC-1991
/ APPLICATION NUMBER: US 07/632,440
/ FILING DATE: 26-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hoerner Jr., Dennis R.
/ REGISTRATION NUMBER: 30,914
/ REFERENCE/DOCKET NUMBER: 38-21(10538)A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (314)537-6099
/ TELEFAX: (314)537-6047
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1377 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
```



```
; MOLECULE TYPE: DNA (genomic)
US-08-553-943-14

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083        Gaps: 0
  Percent Similarity: 85.714  Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-08-553-943-14 ..
Align seg 1/1 to: US-08-553-943-14 from: 1 to: 1377
      2 GluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ::::|||||:|||||:|||||:|||||:|||||:
      1314 GATTGATGGCTGCTTGGAGCTAGATCGAAGTCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-833-485-9

seq_documentation_block:
; Sequence 9, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgette, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-833-485-9

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083        Gaps: 0
  Percent Similarity: 85.714  Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-08-833-485-9 ..
Align seg 1/1 to: US-08-833-485-9 from: 1 to: 1377
      2 GluLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ::::|||||:|||||:|||||:|||||:|||||:
      1314 GATTGATGGCTGCTTGGAGCTAGATCGAAGTCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-137-440-9

seq_documentation_block:
; Sequence 9, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgette, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/137,440
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/833,485
; FILING DATE: 07-APR-1997
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-137-440-9
```

```
alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-09-137-440-9      ..

Align seg 1/1 to: US-09-137-440-9 from: 1 to: 1377

2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||:||||:||||| |||::: |||||
1314 GATTGATGGCTGCTTGGAGCTAAGATCGAAGCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US91-06148A-9

seq_documentation_block:
; Sequence 9, Application PC/TUS9106148A
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1377 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US91-06148A-9

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x PCT-US91-06148A-9      ..
```

```
Align seg 1/1 to: PCT-US91-06148A-9 from: 1 to: 1377

2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||:||||:||||| |||::: |||||
1314 GATTGATGGCTGCTTGGAGCTAAGATCGAAGCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US91-09437-14

seq_documentation_block:
; Sequence 14, Application PC/TUS9109437
; GENERAL INFORMATION:
; APPLICANT: Klee, Harry J.
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: Control of Fruit Ripening and Senescence
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/09437
; FILING DATE: 19911217
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,440
; FILING DATE: 26-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10538)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1377 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US91-09437-14

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x PCT-US91-09437-14      ..

Align seg 1/1 to: PCT-US91-09437-14 from: 1 to: 1377

2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||:||||:||||| |||::: |||||
1314 GATTGATGGCTGCTTGGAGCTAAGATCGAAGCTCCGAC 1355

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-476-008-2

seq_documentation_block:
; Sequence 2, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
```

;; APPLICANT: Kishore, Ganesh M.  
;; APPLICANT: Padgett, Stephen R.  
;; APPLICANT: Stallings, William C.  
;; TITLE OF INVENTION: Glyphosate Tolerant  
;; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
;; STREET: 700 Chesterfield Village Parkway  
;; CITY: St. Louis  
;; STATE: Missouri  
;; COUNTRY: USA  
;; ZIP: 63198  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/476.008  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/306.063  
;; FILING DATE: 13-SEP-1994  
;; APPLICATION NUMBER: US 07/749.611  
;; FILING DATE: 28-AUG-1991  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/576.537  
;; FILING DATE: 31-AUG-1990  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hoerner Jr., Dennis R.  
;; REGISTRATION NUMBER: 30,914  
;; REFERENCE/DOCKET NUMBER: 38-21(10660)A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (314)537-6099  
;; TELEFAX: (314)537-6047  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1982 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 62..1426  
;; US-08-476-008-2

alignment\_scores:  
Quality: 37.00 Length: 14  
Ratio: 3.083 Gaps: 0  
Percent Similarity: 85.714 Percent Identity: 50.000

alignment\_block:  
US-09-485-529-104 x US-08-476-008-2 ..  
  
Align seg 1/1 to: US-08-476-008-2 from: 1 to: 1982  
  
2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15  
:::||||:||||:||||:||||:||||:||||:  
1373 GACCTGATGGCGGGTGGCGCGAAGATCGAAGTCTCCGAT 1414  
  
seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-306-063-2

seq\_documentation\_block:  
; Sequence 2, Application US/083060603  
; Patent No. 5633435  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.

;; APPLICANT: Kishore, Ganesh M.  
;; APPLICANT: Padgett, Stephen R.  
;; APPLICANT: Stallings, William C.  
;; TITLE OF INVENTION: Glyphosate Tolerant  
;; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
;; STREET: 700 Chesterfield Village Parkway  
;; CITY: St. Louis  
;; STATE: Missouri  
;; COUNTRY: USA  
;; ZIP: 63198  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/306.063  
;; FILING DATE: 13-SEP-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/749.611  
;; FILING DATE: 28-AUG-1991  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/576.537  
;; FILING DATE: 31-AUG-1990  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hoerner Jr., Dennis R.  
;; REGISTRATION NUMBER: 30,914  
;; REFERENCE/DOCKET NUMBER: 38-21(10660)A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (314)537-6099  
;; TELEFAX: (314)537-6047  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1982 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 62..1426  
;; US-08-306-063-2

alignment\_scores:  
Quality: 37.00 Length: 14  
Ratio: 3.083 Gaps: 0  
Percent Similarity: 85.714 Percent Identity: 50.000

alignment\_block:  
US-09-485-529-104 x US-08-306-063-2 ..  
  
Align seg 1/1 to: US-08-306-063-2 from: 1 to: 1982  
  
2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15  
:::||||:||||:||||:||||:||||:||||:  
1373 GACCTGATGGCGGGTGGCGCGAAGATCGAAGTCTCCGAT 1414  
  
seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-833-485-2

seq\_documentation\_block:  
; Sequence 2, Application US/08833485  
; Patent No. 5804425  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.

APPLICANT: Stallings, William C.  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1982 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 62..1426  
US-08-833-485-2

alignment\_scores:  
Quality: 37.00 Length: 14  
Ratio: 3.083 Gaps: 0  
Percent Similarity: 85.714 Percent Identity: 50.000  
alignment\_block:  
US-09-485-529-104 x US-08-833-485-2 ..  
Align seg 1/1 to: US-08-833-485-2 from: 1 to: 1982  
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15  
:::|||||:|||||:|||||:|||||:|||||  
1373 GACCTGATGGCGGCGTGGCGGCGAAGATCGAATCTCCGAT 1414  
seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-137-440-2  
seq\_documentation\_block:  
; Sequence 2, Application US/09137440  
; Patent No. 6248876  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/09/137,440  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/833,485  
FILING DATE: 07-APR-1997  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1982 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 62..1426  
US-09-137-440-2  
alignment\_scores:  
Quality: 37.00 Length: 14  
Ratio: 3.083 Gaps: 0  
Percent Similarity: 85.714 Percent Identity: 50.000  
alignment\_block:  
US-09-485-529-104 x US-09-137-440-2 ..  
Align seg 1/1 to: US-09-137-440-2 from: 1 to: 1982  
2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15  
:::|||||:|||||:|||||:|||||:|||||  
1373 GACCTGATGGCGGCGTGGCGGCGAAGATCGAATCTCCGAT 1414  
seq\_name: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:PCT-US91-06148A-2  
seq\_documentation\_block:  
; Sequence 2, Application PCT/TUS9106148A  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.

```
;
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glycosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1982 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 62..1426
;
PCT-US91-06148A-2

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
  Percent Similarity: 85.714      Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x PCT-US91-06148A-2
..
Align seg 1/1 to: PCT-US91-06148A-2 from: 1 to: 1982
2 G1uLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||:||||:||||:||||:||||:||||:
1373 GACCTGATGCCGGCTGGCGCGAAGACTGCACTCCGAT 1414

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-106-761-1

seq_documentation_block:
; Sequence 1, Application US/08106761
; Patent No. 5445956
; GENERAL INFORMATION:
; APPLICANT: HAMMOCK, Bruce D.
; APPLICANT: GRANT, David F.
; APPLICANT: BEETHAM, Jeffrey K.
; TITLE OF INVENTION: RECOMBINANT SOLUBLE EPOXIDE HYDROLASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; STREET: Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/106,761
; FILING DATE: 19930813
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 2307E-445
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1703
;
US-08-106-761-1

alignment_scores:
  Quality: 37.00      Length: 13
  Ratio: 4.111       Gaps: 0
  Percent Similarity: 69.231      Percent Identity: 61.538

alignment_block:
US-09-485-529-104 x US-08-106-761-1
..
Align seg 1/1 to: US-08-106-761-1 from: 1 to: 2101
4 LeuAlaAlaLeuGlyTyrLysValArgAlaSerAspMet 16
|||||  |||||:||||  |||  |||||
879 CTGGCCAGCGAGTTACCGGGTCTAGCTATGGACATG 917

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-629-616-1

seq_documentation_block:
; Sequence 1, Application US/09629616
; Patent No. 6255086
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: OP945CIP
; CURRENT APPLICATION NUMBER: US/09/629,616
; CURRENT FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
```

```
; LENGTH: 4837
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (283)..(1461)
; NAME/KEY: CDS
; LOCATION: (1470)..(4808)
US-09-629-616-1

alignment_scores:
  Quality: 37.00      Length: 12
  Ratio: 3.364       Gaps: 0
Percent Similarity: 91.667 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-09-629-616-1 ..
Align seg 1/1 to: US-09-629-616-1 from: 1 to: 4837
      4 LeuAlaLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15
      |||||:|||||:|||||:|||||:|||||:
4452 CTGGCGTGGTGGCTACAAGATCTCGCCACCGAA 4487

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-182-117-1

seq_documentation_block:
; Sequence 1, Application US/09182117
; Patent No. 6204436
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Transgenic Plant
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,117
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8012 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-182-117-1

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
Percent Similarity: 85.714 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-09-182-117-1 ..
Align seg 1/1 to: US-09-182-117-1 from: 1 to: 8012
      2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ::|||:|||||:|||||:|||||:|||||:
2258 GATTGTGCTGGTGGCTAAGCTGAAGTCTCCGAC 2299

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-182-117-5

seq_documentation_block:
; Sequence 5, Application US/09182117
; Patent No. 6204436
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Transgenic Plant
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,117
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-182-117-4

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
Percent Similarity: 85.714 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-09-182-117-5 ..
Align seg 1/1 to: US-09-182-117-5 from: 1 to: 8418
      2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ::|||:|||||:|||||:|||||:|||||:
2237 GATTGTGCTGGTGGCTAAGCTGAAGTCTCCGAC 2278

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-182-117-4

seq_documentation_block:
; Sequence 4, Application US/09182117
; Patent No. 6204436
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Transgenic Plant
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/182,117
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-182-117-4

alignment_scores:
  Quality: 37.00      Length: 14
  Ratio: 3.083       Gaps: 0
Percent Similarity: 85.714 Percent Identity: 50.000

alignment_block:
US-09-485-529-104 x US-09-182-117-5 ..
Align seg 1/1 to: US-09-182-117-5 from: 1 to: 8418
      2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSerAsp 15
      ::|||:|||||:|||||:|||||:|||||:
2237 GATTGTGCTGGTGGCTAAGCTGAAGTCTCCGAC 2278

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-182-117-4
```

```
Percent Similarity: 85.714   Percent Identity: 50.000
alignment_block:
US-09-485-529-104 x US-09-182-117-4   ..
Align seg 1/1 to: US-09-182-117-4 from: 1 to: 8798
2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSerAsp 15
:::||||:||||:||||:||||:||||:
2347 GATTGATGCTGGTCTTGAGCTAAGATCAAGTCCGAC 2388
seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-472-217-1
seq_documentation_block:
; Sequence 1, Application US/08472217
; Patent No. 5726058
; GENERAL INFORMATION:
; APPLICANT: Alanen-Kurki, Leena
; APPLICANT: Auvinen, Petri
; APPLICANT: Jaakkola, Panu
; APPLICANT: Jalkanen, Markku
; APPLICANT: Lepp, Sirpa
; APPLICANT: Mali, Markku
; APPLICANT: Vihinen, Tapani
; APPLICANT: W rri, Anni
; TITLE OF INVENTION: Syndecan Stimulation Of Cellular
; DIFFERENTIATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/472,217
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,186
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,427
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0050003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(4378..4443, 22026..22106, 23001..23483,
; LOCATION: 23905..24039, 24251..24418)
US-08-472-217-1
```

```
alignment_scores:
Quality: 37.00   Length: 13
Ratio: 3.700   Gaps: 0
Percent Similarity: 76.923   Percent Identity: 61.538
alignment_block:
US-09-485-529-104 x US-08-472-217-1/rev   ..
Align seg 1/1 to reverse of: US-08-472-217-1 from: 1 to: 26700
2 GluLeuLeuAlaLeuGlyTyrLysValArgAlaSer 14
||||||| :|||:||||:||||:||||:
8341 GAATGCTAAGAGTTCAAGCTACCGATGAGTCACT 8303
seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-488-199-5
seq_documentation_block:
; Sequence 5, Application US/08488199
; Patent No. 5851993
; GENERAL INFORMATION:
; APPLICANT: Jalkanen, Markku
; APPLICANT: Mali, Markku
; TITLE OF INVENTION: Suppression of Tumor Cell Growth By
; INTERFERING WITH SYNDECAN-1
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,199
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,862
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0130001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4378..4443
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22026..22107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23002..23483
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23905..24040
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 24252..24418
US-08-488-199-5
```

```
alignment_scores:
  Quality: 37.00      Length: 13
  Ratio: 3.700      Gaps: 0
  Percent Similarity: 76.923      Percent Identity: 61.538

alignment_block:
US-09-485-529-104 x US-08-488-199-5/rev ..

Align seg 1/1 to reverse of: US-08-488-199-5 from: 1 to: 26700

      2 GluLeuLeuAlaAlaLeuGlyTyrLysValArgAlaSer 14
      ||||| : : ||||| : ||||| |||
      8341 GAATTGCTAAGAGTTCAAGGCTACCGAGTGAGGATCAGT 8303
```



